REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL REPORT - 01 Nov 91 - 31 Jan 95

4. TITLE AND SUBTITLE
Fabrication Technology and Measurement of Coupled Quantum Dot Devices

5. FUNDING NUMBERS
61102F
2305/CS

6. AUTHOR(S)
Dimitri A. Antoniadis
Henry I. Smith

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Massachusetts Institute of Technology
Dept of Electrical Engineering and Computer Science
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR- 95-0398

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NE
110 Duncan Avenue Suite B115
Bolling AFB DC 20332-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-92-J-0064

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

SEE FINAL REPORT ABSTRACT

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED

Final Technical Report for AFOSR Contract F49620-92-J-0064

Principal Investigators:

Dimitri A. Antoniadis
Henry I. Smith

Fabrication Technology and Measurement of Coupled Quantum Dot Devices

by

Martin Burkhardt

This report describes the fabrication and measurement of planar tunneling devices. X-ray lithography was used to define gate patterns in order to achieve lateral electrostatic confinement in a two-dimensional electron gas. Technologies were developed for the printing of features with linewidth of 50 nm and below, a lithograpic resolution which is necessary for the fabrication of narrow tunneling barriers. Development of technologies such as this can also be used for large-scale fabrication of silicon and GaAs devices and circuits with critical dimensions of 100 nm and below.

Quantum dots in which the capacitances to the dot were minimized, were fabricated using high resolution lithography. Decreased capacitances to the dot increase the charging energy of a quantum dot, making it possible to observe single electron effects at elevated temperatures. The conductance of a device, featuring eight electrodes to control size and shape of a quantum dot, was measured in a Heliox insertion probe at a temperature of 300 mK. Measurements of several quantum dot sizes were performed and the results were discussed. The same device was biased to produce two unequal quantum dots in series. The results are discussed and compared with theoretical predictions.

Date: March 31, 1995

Contents

List of Figures

Chapter 1

Introduction

Since the invention of the planar silicon process about three decades ago, device dimensions have shrunk dramatically, enabling circuit designers to increase the complexity of integrated circuits. This progress was achieved through novel techniques in device design and layout as well as lithographic technologies. Considering the critical dimensions of present day circuits and the wavelength of light used in printing these circuits, it is clear that lithography will be the single most important factor in the miniaturization of future circuits. Figure 1-1 shows the shrinking of the critical dimension (CD) over the last few years for successive generations of devices, including projections for future generations. Each generation is labelled by its corresponding dynamic random access memory (DRAM) generation. Up to now, DRAM development was the main driving force for advances in lithography, but lithography for logic applications is becoming equally demanding. Requirements on overlay control - the ability to match different mask levels on a substrate - are also getting more stringent with every new generation. Control of the linewidth is typically set at 10 % of the critical dimension. The ability of a lithographic technology, whether it be optical, e-beam or x-ray lithography, to meet these requirements will ultimately decide which one will find its way into the manufacturing environment.

While it is very likely that optical lithography will be sufficient for the 0.25 μm

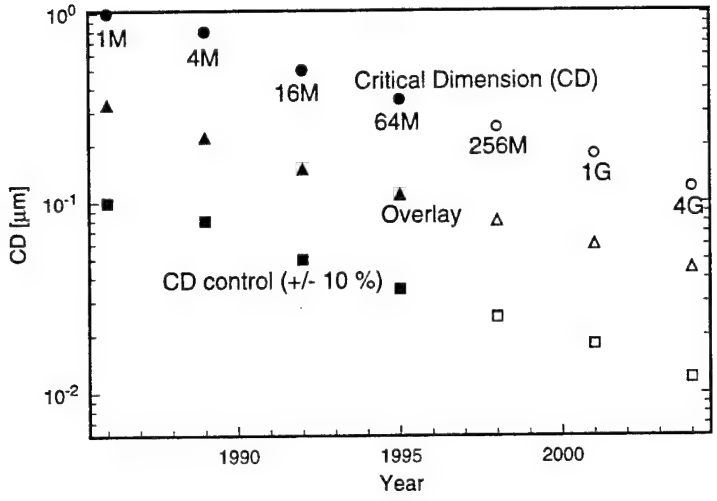


Figure 1-1: Critical dimension (CD) for successive DRAM generations. Critical dimensions follow approximately a straight line in this semi-log plot. Other important lithography parameters, such as overlay tolerance and critical dimension control are parallel to this line (Typical CD-control is 10 % of CD).

technology, which is anticipated to be in production in 1997, the next two generations of 0.18 μm and 0.12 μm will have to be realized with extremely sophisticated optics, using deep UV lasers at KrF- and ArF- wavelengths close to the diffraction limit. At such short wavelengths, absorption within the lens can cause heating, changing the index of refraction, which in turn can move the focal plane. It also becomes increasingly difficult to design lenses with high numerical aperture in the deep UV regime. Furthermore, new resist systems have to be developed for every new wavelength generation (i.e. G-line, I-line, KrF, etc.). The optical lithography community is constantly developing new schemes to enhance the performance of tools for a given wavelength, such as off-axis illumination and phase-shifting masks. Even if it is possible to design such optics with the necessary field size and absence of distortions and aberrations, the requirements for the wafer and underlying relief structure will be extremely strict, requiring expensive processing steps. One of the main problems is the reduced depth of focus for shorter wavelength, which will be well below 0.5 μm and will require extremely flat wafers and polished surfaces. Another problem for optical lithography is

reflection off surfaces and interfaces to be exposed, which will require anti-reflective coatings. It is questionable if a solution which uses deep-UV lithography will still be economically feasible at CDs of 0.18 μm and 0.12 μm. X-ray lithography could be in a strong position as the only available alternative high-throughput, high resolution lithography. While problems concerning overlay and mask production have to be addressed, x-ray lithography does have significant advantages in relief coverage [1], and absence of reflections.

For this reason, it is essential to develop technologies needed for x-ray lithography and to use these techniques for applications that make use of its advantages. In this area, some ground-breaking work has been done at MIT's Nanostructures Laboratory. One of the advantages, superb resolution, has been used to fabricate quantum-effect devices [2–10], high performance silicon devices [11,12], and T-gates for MESFET devices [13]. X-ray lithography has also been used to pattern large area coherent deep submicron patterns [14–16], a task difficult to achieve with other lithography tools except for direct exposure using holographic lithography.

It is also questionable if it will be possible to use old paradigms in the search for future electronics. The shrinking of device dimensions seemed to be limitless for a long time, but in the future it might not be feasible to continue to scale down devices. There are classical issues to be dealt with, e.g. problems arising from the power dissipation of the devices which make it very difficult to cool highly-integrated circuits and more fundamental issues, e.g. the quantization of charge, which makes the movement of small electron clouds between devices a statistically unreliable process. One of the most problematic issues is the so-called wiring crisis in traditional circuit design. Even if one develops technologies to implement extremely high integration densities, one still has to connect all the devices, leading to wiring problems on the chip. One promising solution for these problems is the use of quantum-effect devices[17]. The nonlinear characteristics of quantum-effect devices such as resonant tunneling devices and single-electron transistors could be used for logic applications.

Arrays of quantum-effect devices might be able to do computation by nearest-neighbor interaction, thus offering a solution to the wiring crisis. For this reason, it is essential to investigate the area of mesoscopic physics and try to transfer as much into the engineering community as necessary to create new paradigms.

In this report, I will describe the fabrication and measurement technology that is necessary to build planar tunneling devices which manifest quantum effects. I have chosen an unusual format in that the core of the report is contained in only two chapters, one chapter containing x-ray lithography and device fabrication, and the other device operation and measurement. This division reflects the split of my work between fabrication work, working on x-ray lithography and device fabrication in the NanoStructures Laboratory on the one hand, and measuring the fabricated devices in the measurement laboratory on the other. A further division of each of these chapters seemed artificial and not appropriate. The fabrication chapter contains a short introduction to x-ray lithography, then describes the patterning and electroplating of x-ray masks, and finally details the device fabrication. The measurement chapter begins with an introduction to the problem, briefly reviewing reduced-dimensional systems. It then describes the theory and measurement of a quantum dot device, followed by the theory and measurement of a double quantum dot device. After these core chapters, the work is summarized and put into perspective with general developments in device research, and logical continuations of this research are proposed. In order not to interrupt the train of thought, extensive parts describing details have been transferred to appendices. They describe the fabrication sequence for devices, x-ray mask patterning, the layout transfer from MIT to NRL, diagnostic patterns for e-beam mask writing, an x-ray lithography run-sheet, and electronic circuits used for device measurement, x-ray mask gap control, and electroplating.

Chapter 2

X-Ray Lithography and Device Fabrication

2.1 Introduction

The x-rays typically used for lithography are soft x-rays with wavelengths in the 1-nm region. In our laboratory, we use Cu-L x-rays with a wavelength of 1.3 nm. Refractive optics for use in the x-ray regime do not exist since all materials have approximately the same index of refraction. This is actually a significant advantage of x-ray lithography over optical lithography because of the absence of reflections and standing waves during exposure. The only significant difference between materials at these wavelengths arises in absorption. X-rays can be projected using zone plates or partially reflective multilayer mirrors, but the approach taken by most groups is to go from the farfield into the nearfield regime and use proximity x-ray lithography.

2.1.1 Diffraction and Penumbra

It is well established that the relationship between the maximum mask-to-substrate gap G and minimum feature size w in proximity x-ray lithography is given by the

relationship

$$G = \alpha \frac{w^2}{\lambda} \quad , \tag{2.1}$$

where λ is the x-ray wavelength. In practice, the maximum value for the parameter α is equal to unity [18–20], in some cases 1.5 or even larger [21,22]. For λ=1 nm, G must be $\lesssim 5\,\mu$m for $w < 70$ nm. To achieve controlled gaps below 5 μm requires the development of new technologies such as extremely flat x-ray masks and the means of detecting and removing dust particles [23] that would otherwise interfere with the achievement of such gaps. Experience in other fields (e.g. magnetic recording) indicates that eliminating dust and working with small gaps is challenging but not a fundamental problem.

A large gap can be detrimental to the printing of fine features using proximity x-ray lithography for another reason: if the source has a finite extent, there exists a penumbral blur, δ, of the image cast onto the substrate, which can be calculated if the gap, source extent, and source-to-mask distance are known. With this in mind, one defines a figure of merit β such that [18]

$$\delta = \beta w \quad , \tag{2.2}$$

where w is the minimum feature size to be printed. For a spatially coherent source, e.g. a plane wave, β vanishes. These two figures of merit, α and β, are used to describe the effects of gap on exposure contrast. It turns out that β and α are coupled, that is, for a given w, α increases for increasing β until $\beta \simeq 0.5$. To achieve a maximum value of α (e.g. $1 < \alpha < 1.5$), β should be $\simeq 0.5$.

2.1.2 X-ray Mask Standard

The x-ray mask technology for the fabrication of fine patterns was developed by previous graduate students of the Nanostructures Laboratory [1,24,25] and is described in ref. [26]. The present-day mask standard and most of the work done in our laboratory

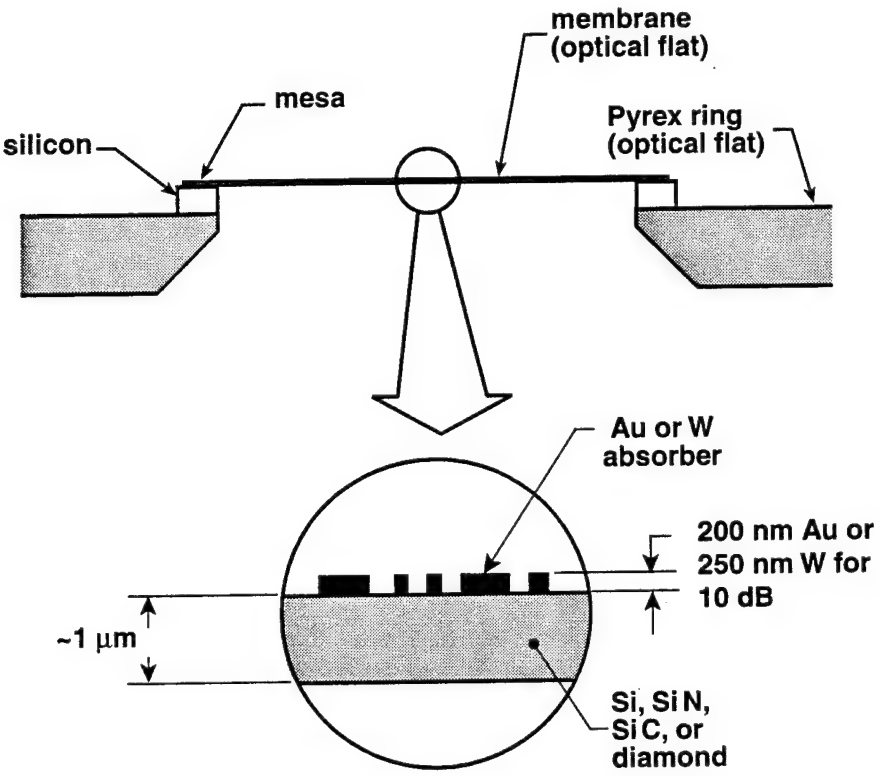


Figure 2-1: Schematic of the mask commonly used for soft x-ray proximity lithography at λ=1.3 nm. In this case, the membrane consists of SiN, the absorber of Au. Other materials are possible, as indicated. Courtesy of Bill Chu.

is based on this technology, with small modifications.

The mask consists of a 31 mm-diameter SiN_x membrane supported on a $\sim 400\,\mu$m-thick Si ring, anodically bonded to a pyrex support frame [26]. A schematic of the mask is shown in Fig. 2-1. Fabrication can be refined to yield masks in which both the membrane and the mesa are optically flat to within less than the wavelength of visible light [27]. A different process, yielding masks with the same geometry, has been shown to produce masks flat to within 500 nm [28]. The Si ring provides an elevation of the membrane above the plane of the pyrex ring of approximately 400 μm. With this "mesa" configuration, the only part of the mask that comes in close proximity to a substrate, or another mask during mask replication, is the narrow mesa rim and the patterned membrane itself.

A plating base, consisting of 5 nm NiCr and 10 nm gold is evaporated onto the membrane to provide a base for the electroplating of gold. Care is taken to cover the edges of the mesa with thick gold to ensure electrical conduction from the pyrex ring onto the membrane. This is accomplished with an additional evaporation of 200 nm of gold outside the membrane.

2.1.3 Absorber Material

The mask blank, described above, is patterned with an absorber of 10 dB attenuation to produce an x-ray mask. The absorber is chosen to have high atomic weight to keep the absorber thickness to a minimum and for ease of patterning. Commonly used absorbers include gold [29–32] and tungsten [25,33,34], used primarily in Europe and the US, and Tantalum, used mainly in Japan. While tungsten and tantalum are patterned using a subtractive dry etching process, gold is patterned via an additive process in which it is electroplated into the interstices of a resist mold. Work described in this report is exclusively done with gold as an absorber, and the patterning and electroplating of this absorber from the mask blank to the patterned mask is described in detail in succeeding sections. The mask patterning process is described in detail in Appendix A.

2.2 Mask Flatness and Particle Control

Ideally, the mesa rim is optically flat [28]. For the older-style masks used in most of the experiments reported here, the mesa rim deviates from a perfect flat by about 1-3 μm. This nonflatness sets an initial minimum gap for experiments in which the mask is brought in proximity to a wafer or another mask. It is not a great concern when the mask is used to print onto substrates that are small compared to the diameter of the membrane. If one copies one mask onto another, the nonflatness of the mesa rim provides an initial gap of 2 to 6 μm. For an ideal, extremely flat mesa rim, studs can

be placed on the rim to establish an initial gap.

To check for the presence of particles and for flatness of an x-ray mask, a glass optical flat with a ~5 μm-thick spacer made of photoresist is used, as shown in Fig. 2-2. Any patterned or unpatterned clear field x-ray mask can be placed onto this optical flat. The nonflatness of the mask can easily be checked in green light by observing the interference fringes caused by the gap between the mask and flat. The gap can be checked by counting the fringes that appear by moving the point of observation from a 90^o angle to a 45^o angle. The gap is then approximately the number of fringes minus one in microns. This gap measurement technique was described by Schattenburg and others [35]. The gap between the mask and flat should be verified to be approximately the thickness of the photoresist. The gap typically turns out to be a little larger than the resist thickness because of the nonflatness of the mesa.

Particles on the mask or flat that are larger than the mask-to-flat gap can easily be detected by observing the fringe pattern in green light. Masks are frequently checked using this technique to detect possible particles on membranes or mesas before x-ray exposures. Figure 2-2 shows a schematic and a photograph of a mask on an optical flat, illuminated with green light.

2.3 Patterning of Masks

2.3.1 Process Outline

A typical sequence for the fabrication of the gate level of a quantum effect device is outlined in Fig. 2-3. The fabrication process evolves from the mother mask (a), as written by electron beam lithography, to the daughter mask (b), to the device (c), which in this figure is a monitor sample consisting of Ti/Au (20/80 nm) on silicon. The mother mask (e.g. Fig. 2-3(a)) is exposed with a 50 keV electron beam using typically between 220 and 300 nm-thick PMMA as a resist. The mask is then electroplated with 200 nm of gold. After the completion of the mother mask, a daughter

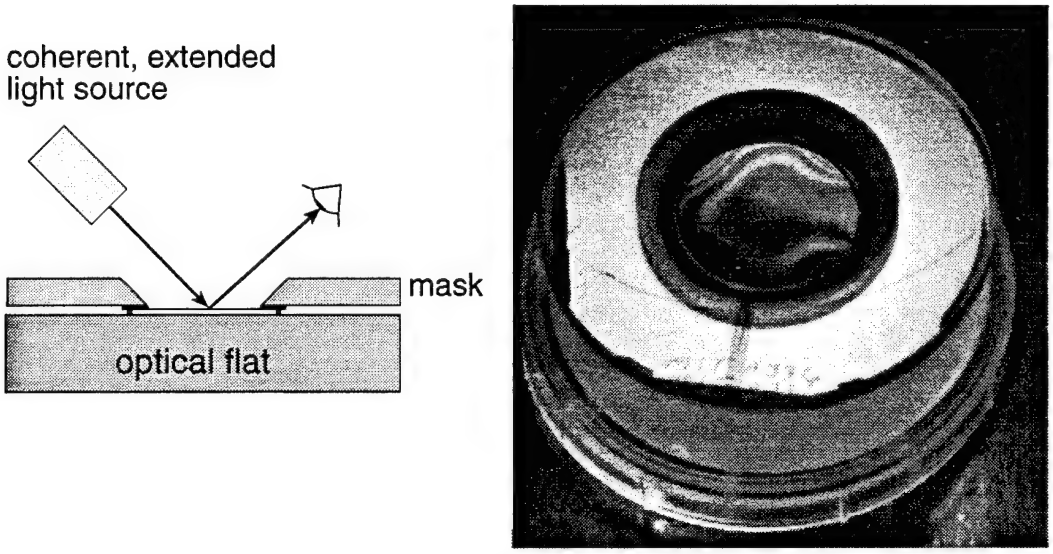


Figure 2-2: Schematic and photograph of particle detection scheme. The x-ray mask is placed on the optical flat to check for particles and flatness.

mask (e.g. Fig. 2-3(b)) is made using x-ray lithography at a gap between 1 and 5 μm. Our x-ray source is a water-cooled electron bombardment target emitting Cu-L x-rays. For masks containing large features (>0.1μm), a gap of about 5 μm may be used. Masks containing very fine features are replicated at a gap below 1.5 μm. Because x-ray exposure is used instead of e-beam exposure, the process latitude in the development (i.e. development time and developer temperature) is much larger. The daughter mask is then used to expose substrates in intimate contact (Fig. 2-3(c)). This intimate contact is achieved by creating a partial vacuum between the substrate and the mask, as described in earlier papers [9,36].

2.3.2 Mother Mask

The process of writing high density patterns with an e-beam is rather time consuming. This makes e-beam lithography unsuitable for large-scale manufacturing. In our case, however, a single x-ray mask is written by e-beam lithography. This mask is then replicated onto a daughter mask, which is then used in x-ray lithography. X-ray lithography is a parallel rather than a serial exposure process (as is e-beam lithogra-

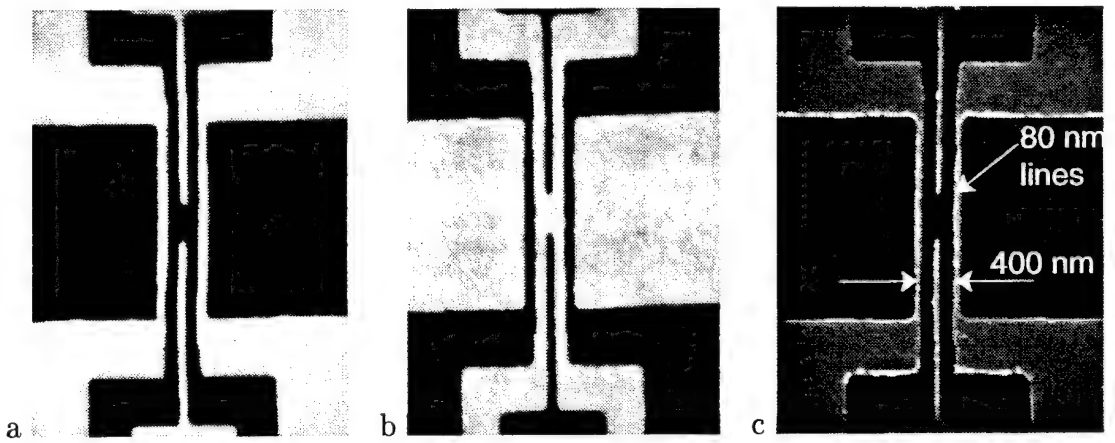


Figure 2-3: a: mother mask (200 nm gold thickness); b: daughter mask (200 nm gold thickness); c: finished gate pattern (20/80 nm Ti/Au) of a device with a point constriction and a tunneling barrier on each side. The center-to-center distance between the gates is 200 nm, the linewidth is about 70 nm.

phy) and, consequently, can have a higher throughput. Our laboratory often writes even large area patterns such as pads and bounding boxes using e-beam lithography. If this is not tolerable, one can utilize a mix-and match scheme to pattern pads with optical lithography and write only fine features by e-beam [27].

After a mask pattern has been designed, it is converted into a data format suitable for a JEOL JBX-5DII electron-beam lithography system. The mother mask is written at the Naval Research Laboratories (NRL) in Washington D.C. The pattern data are transferred electronically to NRL by means of the Internet. The specifics of the transfer process and the patterning have been described in previous theses [24,37], and an updated description is given in Appendix B. The e-beam writing at NRL uses an acceleration voltage of 50 keV and different beam currents for different-sized features. For coarse features ($\sim$10 μm), a beam current of 3 nA and an areal dose of approximately 400 μC/cm^2 is used. For submicron features, a beam current of 15 to 50 pA and an areal dose of 400 μC/cm^2 is typically used. Critical freatures in lateral tunneling devices typically require as narrow a linewidth as possible. For the finest linewidths, those below about 100 nm, a line dose of as low as 1 nC/cm is used. In this case, the electron beam sweeps across the length of the line only once, exposing it

in a single pass. Typically, the same die is written at several locations on a membrane with different areal and line doses assigned to each die to ensure that at least one die satisfies the design with acceptable linewidth results.

After the e-beam exposure, gold is plated up to a thickness of 200 nm, which provides a contrast of approximately 10 dB for the Cu-L line x-rays used in our laboratory. To plate to a thickness of 200 nm, one should use a resist thickness of at least 220 nm. This seems to make the lithographic definition of fine lines more difficult than in a process using direct write e-beam which typically uses thin resist (~50-100 nm). This disadvantage is at least partially compensated by the fact that in the case of an x-ray mask one is writing on a thin membrane rather than a bulk substrate. Most of the electrons are transmitted through the membrane, thus greatly reducing the amount of back-scattered electrons, and therefore the proximity effect compared to bulk substrates [9]. Because of the near absence of a proximity effect, it is possible to achieve linewidths below 50 nm in high density patterns. Figure 2-4 shows such an example, in this case the line is about 30 nm wide. Linewidth as narrow as 25 nm have been achieved this way.

2.3.3 Daughter Mask

After the mother mask has been completed, a copy, the daughter mask, is produced, using x-ray lithography. This replication serves two purposes: it reverses the polarity and produces a dark-field mask suitable for liftoff in device processing; and it reduces the time on the e-beam lithography tool. If a mask should break, a new daughter mask can be made within a short amount of time.

The x-ray exposure for mask replication is done with a microgap between mother and daughter mask. The size of the maximum practical gap is determined by feature size using eq. 2.1, with α set equal to 1 or 1.5. The masks are placed on top of another, with the gap set by either studs evaporated onto the mesa or by the natural nonflatness of the mesa. In both cases, the gap is targeted between 2 μm and 6 μm.

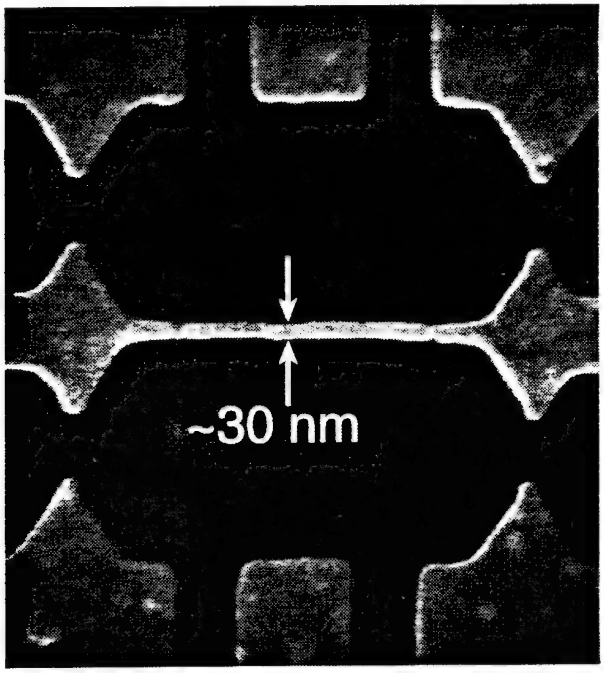


Figure 2-4: SEM micrograph of a device written directly with an electron beam on a mother mask. The line is approximately 30 nm wide.

For masks with ideally flat mesa rims, one can put studs on the rim of the mother mask to establish an initial gap. The gap can then be reduced to about 1 μm by partially evacuating the space between the masks [38].

In the following paragraphs, the calculation of gap and gap capacitance is demonstrated and the apparatus for replicating x-ray masks at a controlled gap of $\sim 1\,\mu$m is described [38]. Such a gap is necessary for faithful replication of mask features $\lesssim 40$ nm.

A schematic of the microgap mask replication apparatus is shown in Fig. 2-5. The capacitance between the two masks is sensed using a Boonton 72B capacitance meter and fed back into a circuit that adjusts the partial pressure between the masks via an MKS 248A control valve to keep the capacitance constant. The pressure between the masks reaches a constant value within a few seconds.

In order to calibrate the apparatus, gaps were measured both capacitively and interferometrically. The latter is easily accomplished using a monochromatic source

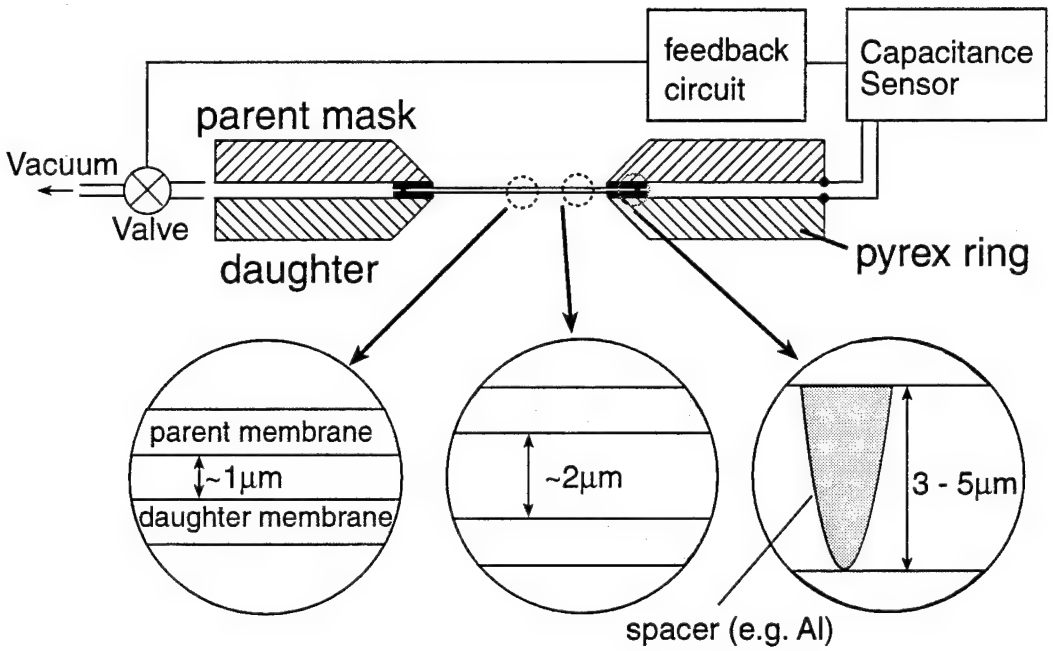


Figure 2-5: Schematic of microgap mask replication apparatus. In this case, the mother mask features studs, in order to keep the masks at a certain gap across the mesa. This technique is necessary for optically flat mesas, while a nonflat mesa provides for a natural gap of about the same size. A Boonton 72B bridge was used as capacitance sensor, an MKS 248A as control valve.

while viewing the fringes between the two mask membranes [35]. Before inserting the masks into the apparatus, any large dust particles should be removed. This can be done using an optical flat, as described earlier.

Calculation of membrane deflection

The gap can be measured and monitored during the exposure setup with the monochromatic green light. The gap can also be calculated using elementary theory of elasticity. In the case of two circular membranes, one can use polar coordinates to analytically solve for the bow of the membranes. The calculation for a similar problem is done in Ref. [39]. If the pressure difference between the top and the bottom of the membrane is denoted as P, we have to solve the following equation

$$\frac{1}{r}\frac{d}{dr}(r\frac{d\zeta}{dr}) = -\frac{P}{T} \ , \tag{2.3}$$

where T is the absolute magnitude of the stretching force per unit length of the edge of the membrane, ζ is the deflection, and r is the radius variable as indicated in the inset of Fig. 2-6. The solution for the deflection of the membrane is

$$\zeta = \frac{PR^2}{4T}(1 - (\frac{r}{R})^2) \ , \tag{2.4}$$

where R is the radius of the membrane. Note that only the amplitude, but not the shape, of the deflection depends on parameters like pressure and forces on the membrane edges. The deflection can also be measured for the gap between the two masks. If the masks have identical physical properties, the deflection of a single membrane should be half the gap between the masks. The gap as a function of radius variable r can be found by determining the location of the interference fringes. The measured deflection in Fig. 2-6 is in good agreement with the fit using eq. (2.4).

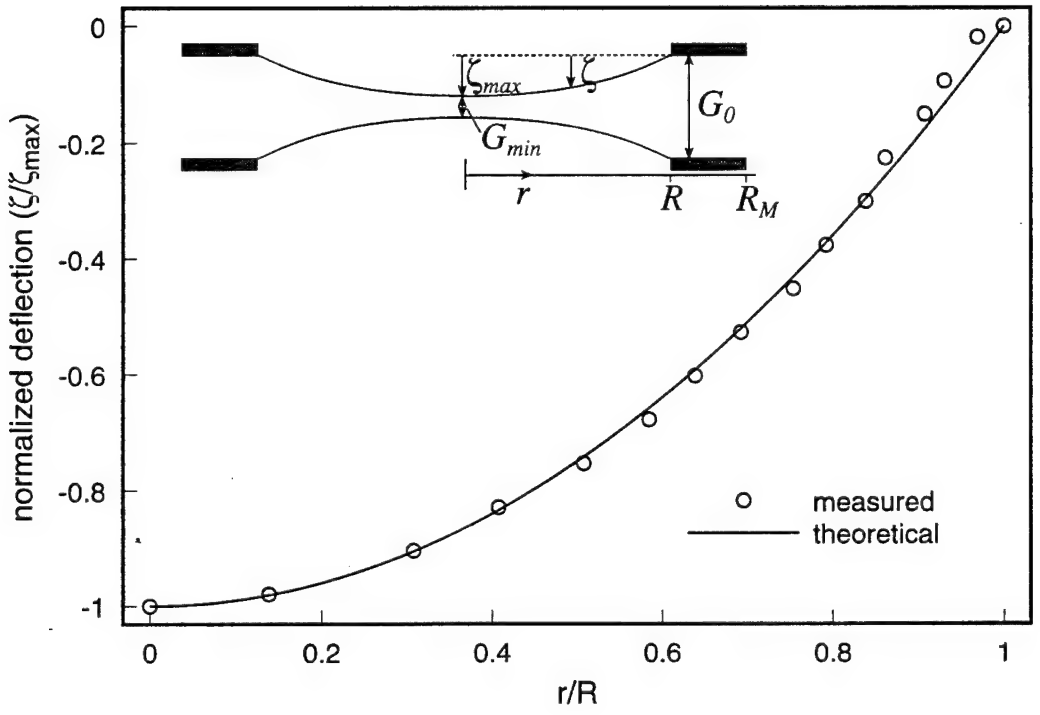


Figure 2-6: Plot of normalized deflection as a function of normalized radius. Measured values are indicated with circles, the line shows the fit using eq. (2.4). The inset shows a schematic of the system, indicating the variables used in the calculations.

Calculation of capacitance vs. gap

Given eq. (2.4), one can calculate the capacitance as a function of minimum gap. The capacitance C' of the membrane is defined as:

$$C' = \epsilon \int_0^{2\pi} \int_0^R \frac{r dr d\theta}{G_0 - 2\zeta(r)} \tag{2.5}$$

where ϵ is the dielectric constant and G_0 is the gap without any force applied, which is taken to be uniform on the circumference of the membrane. The deflection is independent of θ, so

$$C' = 2\pi\epsilon \int_0^R \frac{r dr}{G_0 - \frac{P}{2T}(R^2 - r^2)} \ . \tag{2.6}$$

We observe that the maximum deflection of each membrane is related to the minimum gap G_{min} and G_0 as

$$\zeta_{max} = \zeta(r = 0) = \frac{PR^2}{4T} = \frac{G_0 - G_{min}}{2} \ . \tag{2.7}$$

This leads to

$$C' = \frac{2\pi\epsilon R^2}{G_0 - G_{min}} \int_0^R \frac{\frac{r}{R} d\frac{r}{R}}{\frac{G_0}{G_0 - G_{min}} - 1 + (\frac{r}{R})^2} \ . \tag{2.8}$$

Now we can substitute $X = G_0/(G_0 - G_{min}) - 1 + (r/R)^2$, so that $dX = 2(r/R)d(r/R)$. The boundaries of the integral are $G_0/(G_0 - G_{min}) - 1$ and $G_0/(G_0 - G_{min})$, and we get

$$C' = \frac{\pi\epsilon R^2}{G_0 - G_{min}} \ln\left(\frac{\frac{G_0}{G_0 - G_{min}}}{\frac{G_0}{G_0 - G_{min}} - 1}\right) \ . \tag{2.9}$$

This is the capacitance due to the deflected membrane. There is also a capacitance due to the mesa which does not deflect. The total capacitance is then

$$C = \frac{\pi\epsilon R^2}{G_0 - G_{min}} \ln\left(\frac{G_0}{G_{min}}\right) + \frac{\pi\epsilon(R_M^2 - R^2)}{G_0} \ . \tag{2.10}$$

Here R_M is the radius of the mesa, which in this case is 3 mm larger than the radius of the membrane. The capacitance C as a function of minimum gap G_{min} is plotted in Fig. 2-7 for a range of initial gaps G_0 between 2.5 and 6 μm. These values were calculated using eq. (2.10) and are plotted as full lines. The minimum gap as a function of capacitance was also measured by determining the initial gap G_0 and then counting fringes for different values of the capacitance. Those values are plotted as circles in Fig. 2-7. The measured values are offset from the theoretically predicted values by about 0.8 nF, which can be explained by an uncertainty in the calculation of C_{mesa}.

Scanning electron micrographs of a pattern on the parent x-ray mask made by e-beam lithography and its replication onto a daughter mask using the microgap mask replication apparatus at a gap of $\sim$1 μm are shown in Fig. 2-8. This pattern is used to define the gates in a Coulomb-blockade device that can be biased to yield two quantum dots, coupled to each other via a narrow tunneling barrier, or as one quantum dot with three leads [40,41]. The lithographic barrier width is less than 50 nm, which is critical for the functioning of the device, i.e. for efficient tunneling. While the tunneling barrier - the line dividing the two dots - printed without loss of resolution, the corners appear slightly rounded. Since the gold is plated almost to the thickness of the resist, a top view of a plated mask accentuates any rounding of edges on the top of the photoresist.

2.4 Development

After e-beam exposure, the mother masks are developed in a mixture of methyl-iso-butyl-kytene (MIBK) and iso-propyl-alcohol (IPA). Mother masks are generally immersion developed because it is easier to maintain controlled and repeatable conditions in a beaker than by puddle development. The main disadvantage of immersion development is that the fluid is not constantly being replenished and particles floating

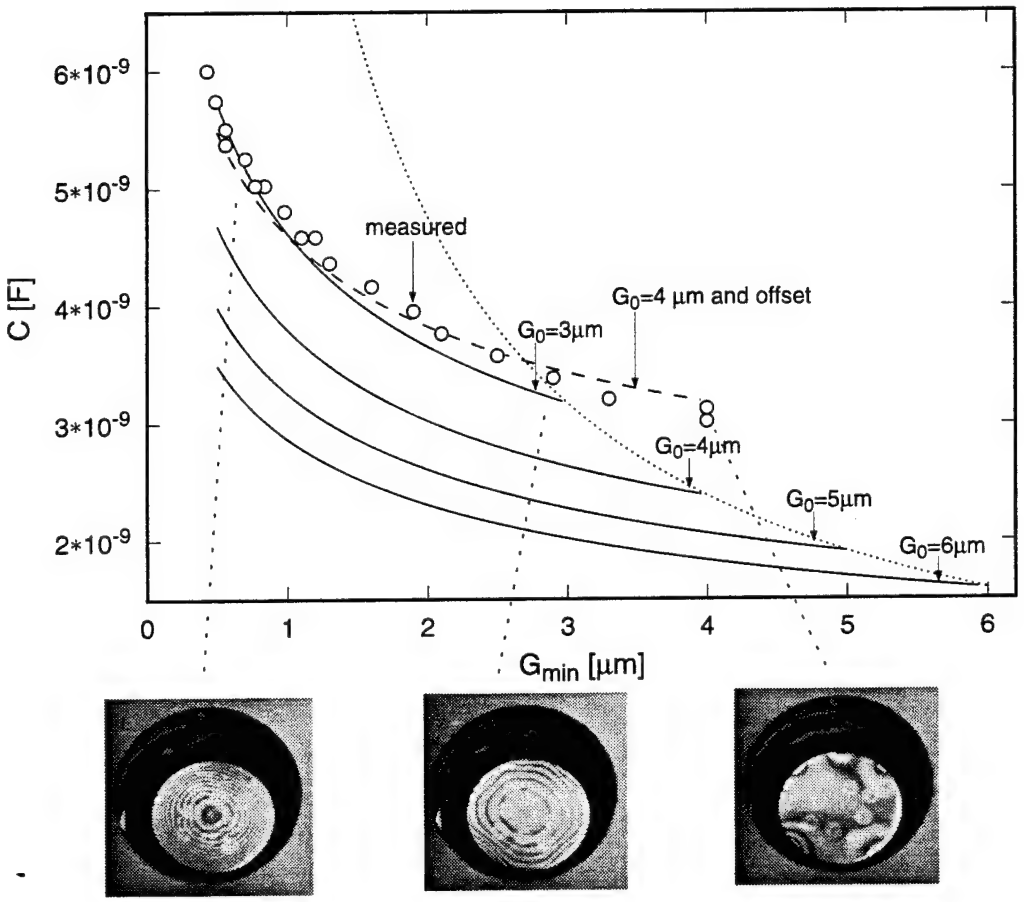


Figure 2-7: Plot of capacitance versus minimum gap. The full lines correspond to theoretical capacitances for different initial gaps between 3 and 6 μm. The measured points are determined by measuring the initial gap and then counting interference fringes. The dashed line is the capacitance for a 4 μm initial gap with an additional offset accounting for uncertainty in C_{mesa}. The dotted line gives the capacitance of a parallel plate capacitor for comparison. The images on the bottom are photographs of the setup at different values of minimum gap.

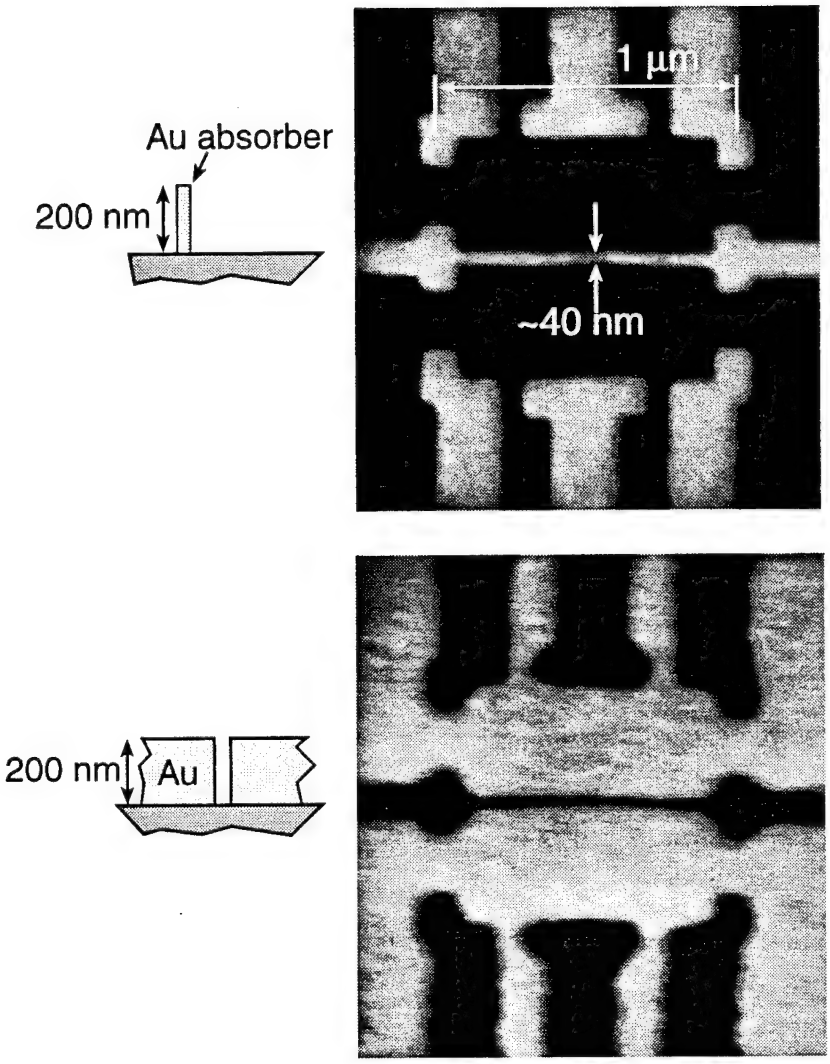


Figure 2-8: Scanning electron micrographs of parent and daughter masks of a pattern containing 40 nm features. The gold absorber is plated to a thickness of 200 nm.

on the surface of the developer can attach themselves to the sample. The development is usually done in a 2:1 IPA:MIBK solution at a temperature of 21 °C for 90 seconds. Upon mixing of the two chemicals, the temperature drops to about 16 °C. The temperature has to be increased through heating and stabilized at the desired point.

Daughter masks, which are x-ray exposed, have a larger exposure latitude than e-beam exposed mother masks and are therefore generally puddle developed with a mixture of 3:2 IPA:MIBK. Puddle development means that a squirt bottle is used to constantly supply fluid to a puddle on the sample, thereby constantly replenishing the fluid. This method has an advantage in that reduces particle contamination problems.

All masks are subjected to a de-scum process between development and plating. PMMA is believed to leave a very thin (mono) layer of scum on the developed surface which would cause poor electroplating if it were not removed. The parameters are given in Appendix A.

2.5 Electroplating

After development, the masks are electroplated with gold absorbers to a thickness of 200 nm, which corresponds to 10 dB attenuation for the Cu-L x-ray sources that are used in our laboratory ($\lambda = 1.32$ nm).

SEL-REX BDT510, a commercially available gold plating bath, is used for electroplating the absorbers on the x-ray masks. Although the plating solution and its brightening agent are proprietary, some information can be gained from Material Safety Data Sheets. The gold is contained in the solution in the form of gold sulphite. This is a toxic substance, but it is still safer to work with than many competing gold plating solutions, which are typically cyanide-based. If the bath is not contaminated by foreign substances, the purity of the deposited gold reaches 99.9 %, according to the manufacturer.

The quality of the plating must be monitored continuously, and quality indicators such as plating roughness and film stress must be correlated to parameters such as current density, reference electrode voltage, bath temperature, brightener concentration, and agitation. The brightener is a proprietary additive to the plating bath [42], which induces grain nucleation and therefore tends to make the gold fine-grained and "bright". The brightener works by bonding to the gold grain, terminating its growth, thus inducing nucleation of new grains. If the grains are small compared to the wavelength of visible light, there is no diffuse scattering on the surface of the gold, making it appear "bright". Since the chemical composition of the brightener is not known, it is important to check the characteristics of the plating bath periodically.

One important characteristic for a plating bath is the current-voltage (IV) plot. The IV-characteristics of BDT510 change as a function of brightener concentration, as indicated in Fig. 2-9. For reference, typical plating currents used in our laboratory are between 10 and 20 mA for the plating geometry used in these experiments. When a new bath is used, such an IV-measurement can be done after the addition of brightener and periodically thereafter to keep track of the electrical changes that occur during the lifetime of the bath. A typical change of IV-characteristic with time for a bath used in our laboratory is depicted in Fig. 2-10. The addition of brightener significantly reduces the voltages and decreases the slope for the currents of interest in the linear portion of the IV-curve. Over time, the voltage increases for identical current density, and is larger than in the bath without brightener after a couple of weeks. It is obvious that the bath changes its chemistry during its lifetime, possibly due to the oxidation or evaporation of brightener and water. After addition of more brightener, the plating voltage is again reduced, as shown in Fig. 2-10. The characteristics indicated here are for a bath which was not hermetically sealed when not in use and thus allowed for significant evaporation and surface reactions such as oxidation. The bath is now sealed after each use to reduce such rapid deterioration of bath quality, which has caused its average lifetime to apporoximately double.

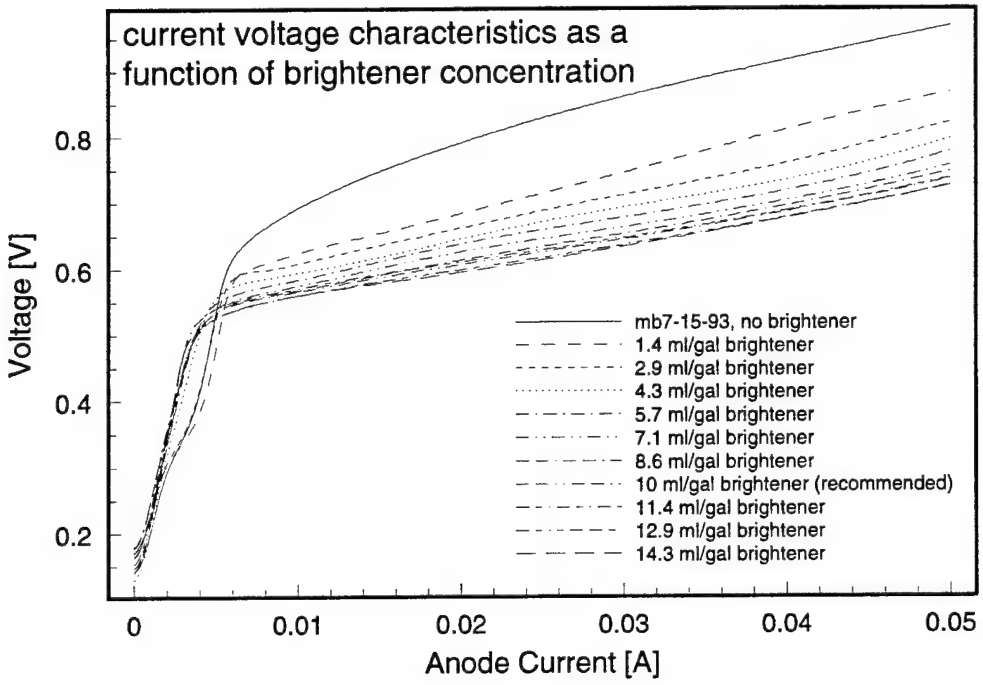


Figure 2-9: Current-voltage-characteristic of a plating bath as a function of brightener concentration. The plot shows the voltage of the cathode with respect to a saturated Calomel electrode. The recommended brightener concentration is 10 ml per gallon. The cathode area in this experiment was about 25 cm^2, so that typical current densities used during plating correspond to between 10 and 20 mA in this plot.

2.5.1 The Plating Apparatus

In order to plate an x-ray mask, a plating apparatus has to be designed and built. Plating in our laboratory is typically done in one large bath which is used for several months and which is strictly monitored during its lifetime [29,24]. Although efforts were made to reduce fluctuations in plating quality by monitoring the pH value, conductivity, stress of the plated film, and temperature, plating was at times unreliable. Problems were usually solved by changing the bath, which consists of about 5 liters of plating solution. In order to reduce costs and to have more flexibility with experiments, a new fixture was designed with the explicit intent of reducing the amount of plating fluid used. A schematic of the fixture is shown in Fig. 2-11. A small beaker,

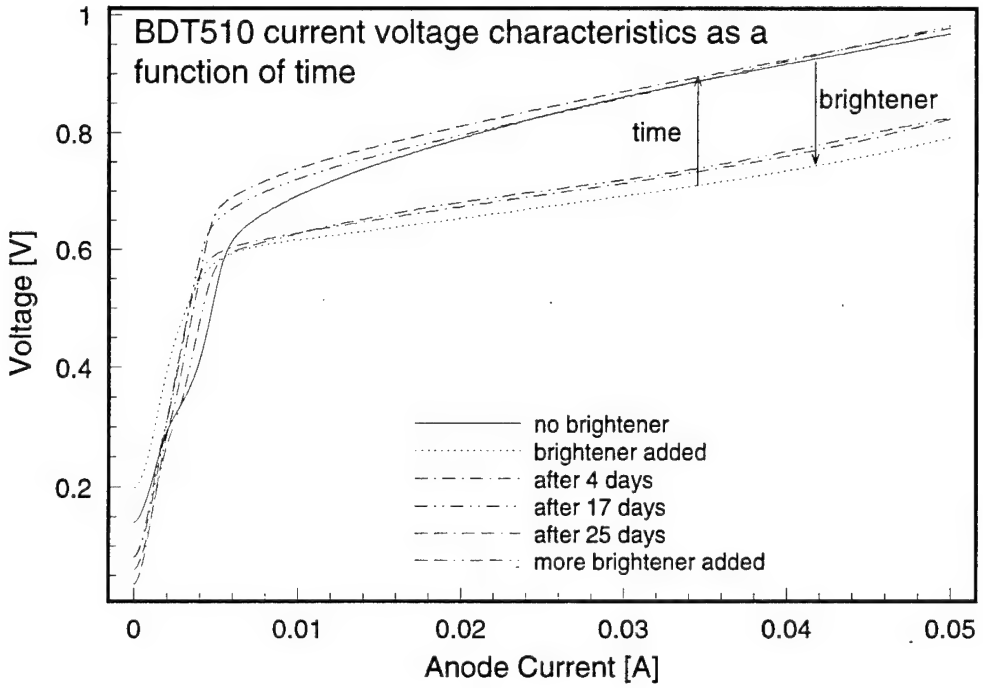


Figure 2-10: Current-voltage characteristic of a plating bath as a function of time. Also shown is a curve for a bath with no brightener. The addition of brightener reduces voltages, whereas they typically increase during the life of a bath. Typical currents, corresponding to current densities used in the fabrication of x-ray masks, are 10 to 20 mA.

containing approximately 250 ml of plating solution, is used. The plating solution is heated on a hotplate to a temperature of 33 °C. The mask is then placed, membrane-side-up, into the beaker, followed by a teflon ring, which serves as a spacer between mask and platinum anode. Gold coated magnet wire on the bottom of the teflon ring ensures electrical contact to the mask, which is the cathode in the plating setup. The insulator of the magnet wire is removed and the copper is gold plated for contact to the mask. The insulator of the magnet wire is effective in the solution. A platinum covered mesh, the anode, is then placed on top of the teflon ring.

One of the drawbacks of this setup is that the solution is not constantly filtered, as is the case in previous setups. In my experience, this has not led to mask defects due to the plating process. In most cases, particles get introduced during fabrication and

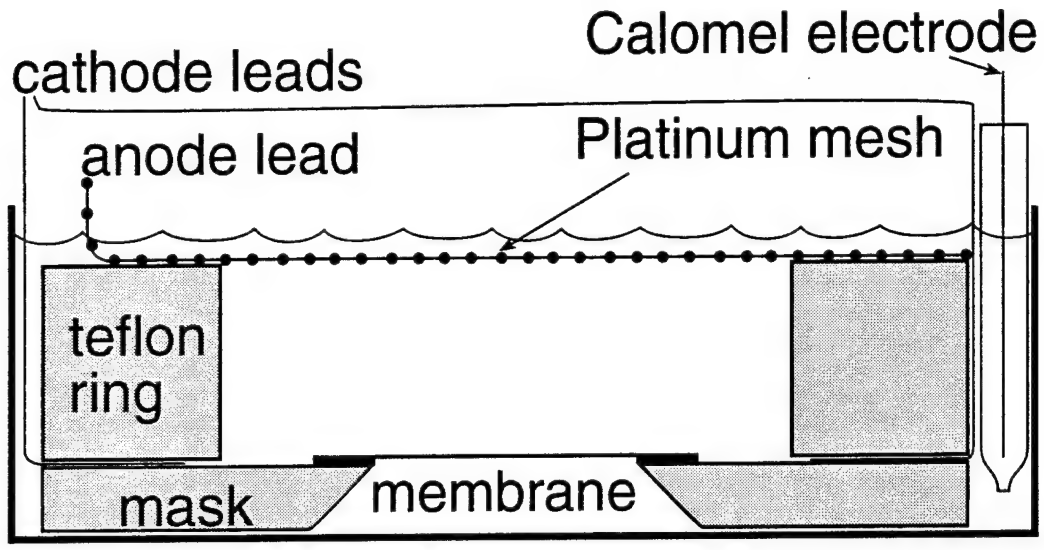


Figure 2-11: Schematic of the fixture used for plating of x-ray masks. A pyrex beaker is used to contain the plating solution, limiting the amount of plating solution to about 250 ml.

handling of the mask blank rather than subsequent processing and blanks can easily be screened for particles before use. The major advantage of the plating setup is that the plating solution can be changed much more frequently, such as once a month or more, without excessive costs. This setup was assembled entirely from components available in the laboratory. A more sophisticated setup, incorporating these ideas, could easily be built.

A power supply, which allows either galvanostatic or potentiostatic control, was designed to be used with the experimental plating setup. In most experiments, the power supply was used as a galvanostat in which the voltage is regulated to impose a constant current. The circuit diagram of this supply and a short explanation of the circuit is given in Appendix C.

If this power supply is used for plating, the value of the area to be plated is set at the plating box. This enables the electronic circuit to estimate a thickness of the plated gold assuming that the plating efficiency is unity. The current is integrated over time and a voltage proportional to the plating current and the plated thickness

is constantly monitored using voltmeters. It is also a good idea to connect a chart recorder to the plating setup, recording simultaneously the plating current and the voltage of the saturated Calomel electrode (SCE). In this way, it is easy to detect problems, even after the plating run is done. The SCE voltage should be similar for comparable samples and equal cathode current densities. Any deviations suggest possible problems with gold quality and bath deterioration.

2.5.2 Determination of Plated Thickness

It is very important to determine the thickness of the plated gold with great accuracy. One generally tries to keep the resist thickness to a minimum in order to avoid problems associated with large aspect ratio resist lines, e.g. falling-over lines, adhesion problems, or line broadening in the e-beam lithography due to forward scattering of electrons in the thick resist. If one wants to plate absorbers to a thickness of 200 nm, the resist thickness used for the e-beam or x-ray exposure is about 220 nm, leaving little room for overplating. It is therefore crucial to end the plating within a few tens of nanometers. If the exact area of the plated surface and the plating efficiency is known, it is possible to plate to within these tolerances using proper timing. The plating efficiency indicates how efficiently the ionic current in the solution transfers metal onto the sample.

The most straightforward way to measure the plated thickness is to use a profilometer, a stylus measurement system such as an Alpha-Step. This has been done on membranes and gives very consistent results. The main disadvantage of this method is the fact that the force exerted by the stylus is not very well controlled and can in some instances destroy the membrane. To avoid such a scenario, a resist step on the mesa of the x-ray mask can be measured instead. The mesa, however, is sometimes far removed from the actual region of interest and there are instances in which the plating is nonuniform across the x-ray mask. This is especially problematic if fine features plate slower than large features, which was occasionally observed in old plating

bath's. In these cases, the voltages observed during plating tended to be higher than normal. In such situations, a profilometer measurement on the mesa might indicate sufficient absorber thickness, while in the regions of interest, towards the center of the membrane, the absorbers are too thin.

A preferred method to measure the thickness of gold plated onto an x-ray mask is to use a microscope and measure the transmission of visible light through the plated region, as indicated in Fig. 2-12. A substage condenser lens is used to image a rectangular aperture onto the membrane at the position of the absorber. Gold is partially transmissive in the visible region, with a local maximum in the green part of the spectrum, and it is possible to relate the intensity of the transmitted green light to a given gold absorber thickness. The substage condenser lens is movable, so that the absorber, which in most cases consists of rectangles, can be imaged by adjusting the aperture to fit the desired field of view. The transmission of the green light through the sample can then be compared to the transmission of the green light through a known thickness of gold. A prerequisite is a similar transmission through the SiN membrane on sample and reference membrane. An advantage of this method is that the lens chosen to observe the thickness through transmission can be a high magnification, high numerical aperture lens, enabling the observer to recognize variations of absorber thickness on a lateral scale of a few microns. Using a 150 W Xenon light source, it is possible to visually detect transmitted green light for gold thicknesses up to about 220 nm for a 10x, 0.2NA lens, and up to just below 200 nm for a 100x, 0.9NA lens.

It turns out that the sensitivity of the human eye, especially at low intensities, makes thickness determination within a hundred Ångstroms possible. It should be noted, however, that this method is not quite safe for the human eye, as a mask breakage would result in a high light intensity incident on the observer's eye. A drawback of the current system is the fact that comparison is possible only after exchanging the sample mask with the mask containing the standard gold thicknesses. This could be

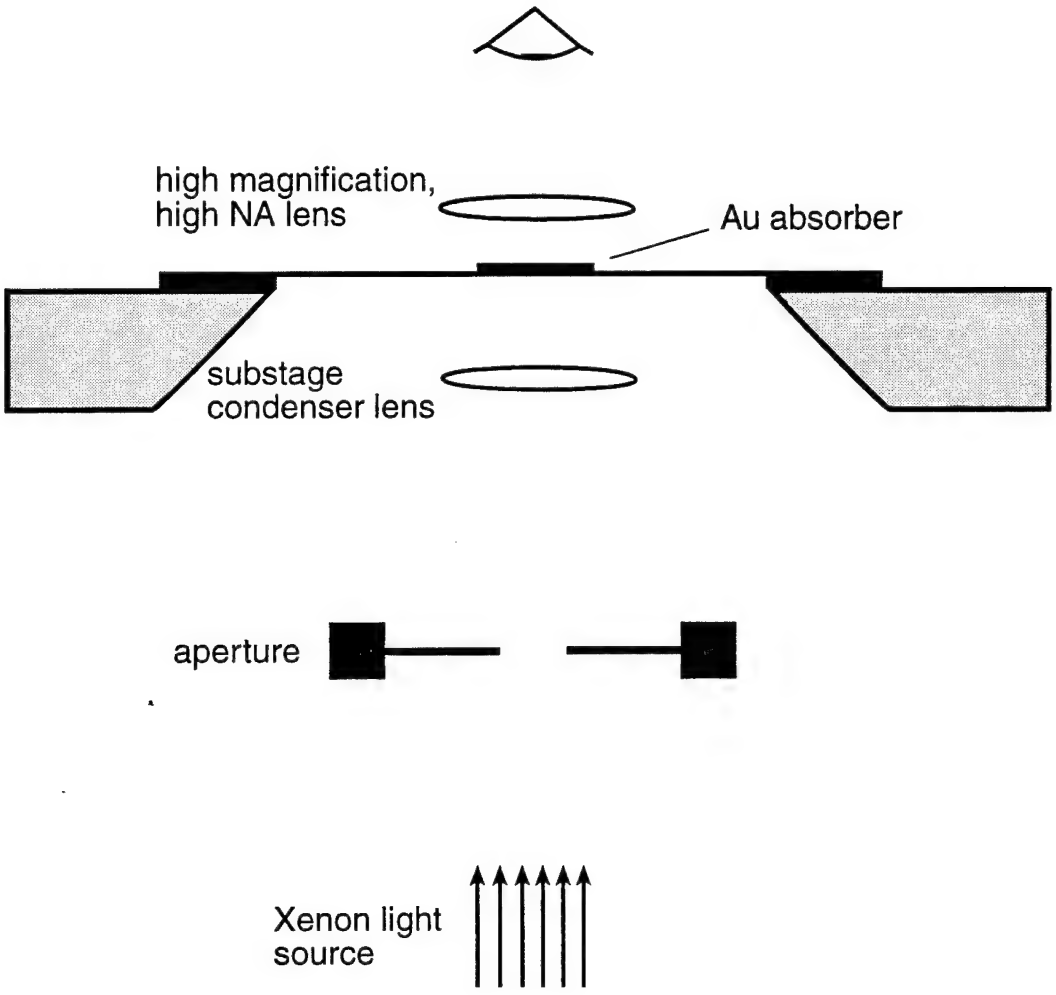


Figure 2-12: Schematic of method used to determine plated gold thickness. A rectangular aperture, illuminated with bright light, is focused onto the backside of a gold absorber. The intensity of the transmitted light is dependent on the gold thickness. A high magnification, high NA lens can be used for observation, permitting high lateral resolution in determining the absorber thickness.

improved easily by introducing glass slides, evaporated with standard thicknesses of 1000, 500, 200 and 100 Ångstroms for example, plus plating base, into the illumination path of the substage condenser. In this way, transparent areas of the mask would be dimmed and one would have to match up this intensity with the intensity of the absorber region without the additional filter present. This method would eliminate the uncertainty regarding SiN thickness of sample and reference membrane and the accuracy of the thickness reading could be further increased.

2.6 Device Fabrication

To make quantum effect devices, we start off with substrates, grown using molecular beam epitaxy (MBE). The layer sequence is similar to layer structures used to build High Electron Mobility Transistors (HEMTs), with the difference that the two dimensional electron gas (2DEG) is closer to the surface. For a comprehensive review on such structures, see ref. [43]. Since constrictions of the electron path are achieved through electrostatic confinement, the electron gas needs to be close to the gate electrode. A typical layer structure used in the experiments is depicted in Fig. 2-13. The material is grown by Prof. Michael Melloch of Purdue University. Starting with a semi-insulating wafer of GaAs, a 1 μm-thick layer of undoped GaAs is grown. On top of that, two layers of $Al_{0.3}Ga_{0.7}As$ are grown, a 20 nm-thick undoped layer and a 30 nm-thick silicon-doped layer. The latter provides the charge carriers that accumulate in the 2DEG in the GaAs close to the AlGaAs interface. The layer structure is terminated by a 5 nm layer of n-doped GaAs. The undoped AlGaAs layer provides a spatial separation between the ionized impurities and the charge carriers, preventing extensive scattering at low temperatures. Typically, the thicker the spacer layer, the larger the mobility and the smaller the charge density in the well. Since we need good control over the potential in the 2DEG and a short distance to the surface, the layer structures used in our experiments represent a trade-off between high mobility, high

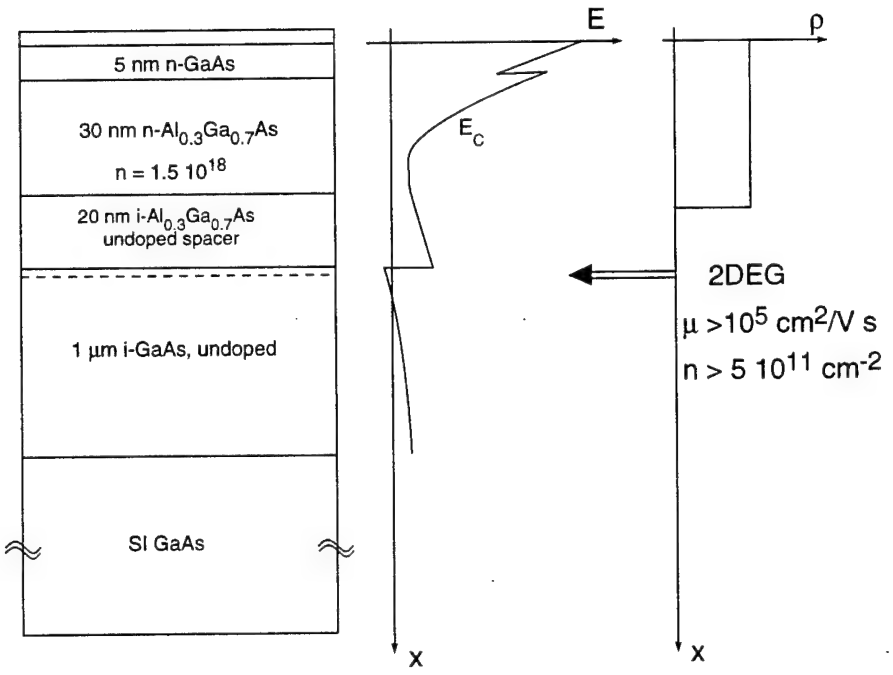


Figure 2-13: Layer structure used as substrate for the fabrication of quantum effect devices.

sheet density, and thin layer structure. Typically, the layers used in our experiments have a mobility of about 250,000 cm^2/Vs and a carrier density of about 5×10^{11} cm^{-2}.

An outline of the fabrication process is shown in Fig. 2-14. Starting with the layer structure, photolithography is performed using an optical mask which defines the mesa structure. The active and the ohmic-contact regions of the devices are defined and device isolation is achieved. I use an etch consisting of 1000:20:6 H_2O:NH_4OH:H_2O_2 for about 20 s, which results in a mesa of about 60 nm.

To put down the ohmic contacts, photolithography is performed using the ohmic contact mask, followed by a short de-scum in a UV-ozone cleaning station and the removal of oxide from the sample surface using a dip in buffered ammonium hydroxide solution. Then the ohmic contact material is deposited by electron beam evaporation. The layer sequence for this evaporation is Ni(2):Au(4):Ge(22):Au(44)Ni(10):Au(30), where thicknesses, indicated in parenthesis, are in nanometers. The contacts are subsequently annealed in a rapid thermal annealing system for 15 sec at 430°C.

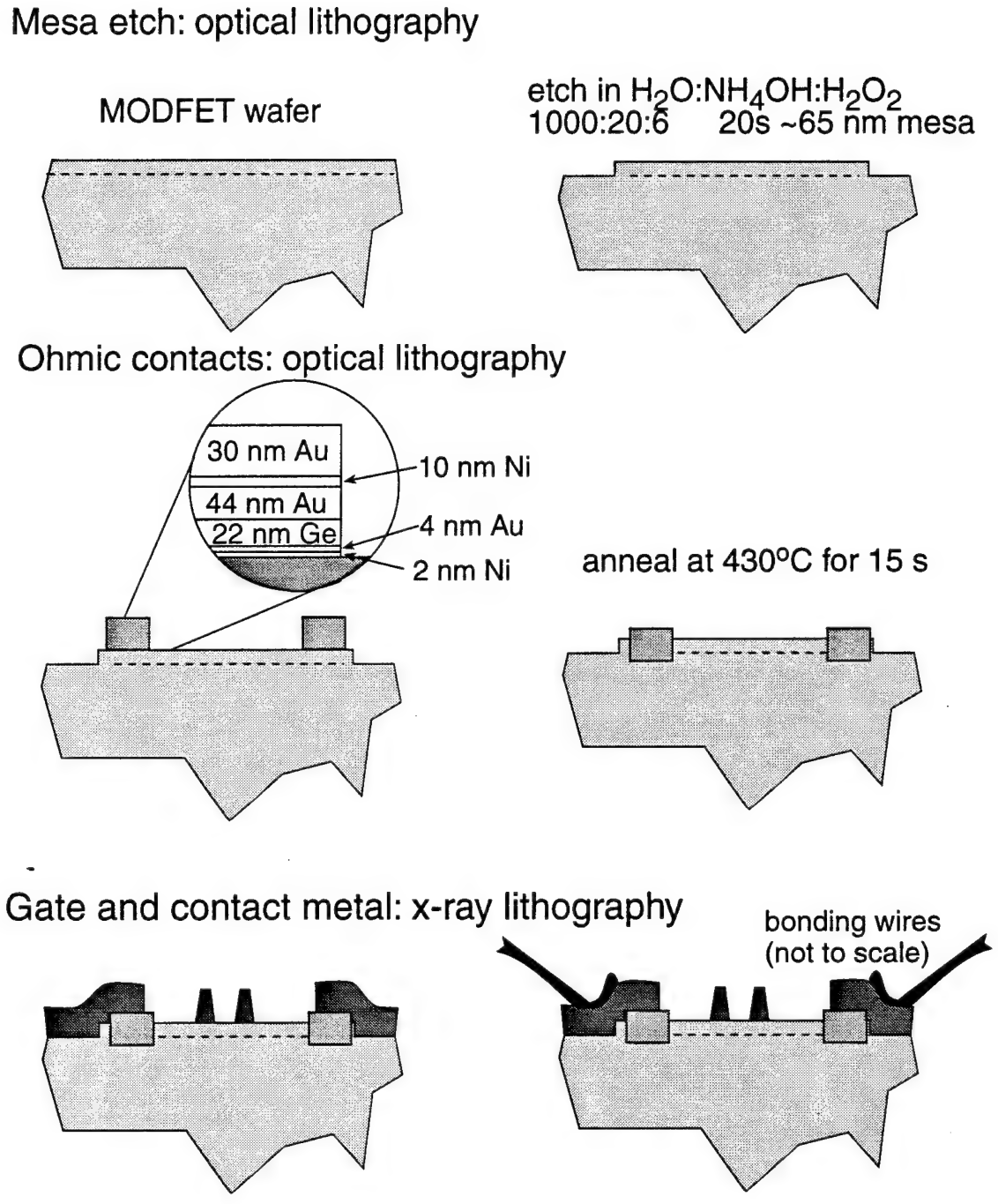


Figure 2-14: Outline of the fabrication process for quantum effect devices.

The last step involves the deposition of the gates which define the geometry of the potential barriers of the quantum effect device. PMMA (950K, 4%) is spun onto the substrate to a thickness of about 250 nm and baked for an hour at 180 °C. The sample is then aligned and exposed following the procedure developed by Ghanbari and Chu [9,36]. The exposure is done in our laboratory using x-ray lithography. After development, about 100 nm of metal are deposited and lifted off. For electron beam evaporators, 10 nm titanium, followed by 90 nm gold, are evaporated. For thermal evaporators, 100 nm of aluminum can be used. There is some indication that use of electron beam evaporators cause damage in 2DEGs [44]. Metals as thick as 120 nm are routinely lifted off using a resist thickness of about 250 nm. The complete fabrication sequence is listed in more detail in Appendix D.

2.6.1 Advantages of X-ray Lithography in Fabrication

There are only a few lithography technologies that can achieve the resolution sufficient for patterning the gate structures of quantum effect devices. Optical lithography has made great progress in achieving small feature sizes but it is still far from the necessary linewidth. Electron-beam lithography is the tool most often used in the fabrication of quantum-effect devices. There is, however, some indication that under certain conditions damage is inflicted on devices during exposure with an electron-beam [45]. In the case of ion-beam lithography, the problem of damage is even more severe, as indicated by Ismail [4]. The only lithography technique that can directly compete with e-beam lithography is x-ray lithography which has been shown to be capable of resolving features down to 10's of nanometers [1], and has further been shown to cause no damage in modulation doped structures [45].

X-ray lithography has some unique advantages over e-beam lithography. Among them is the fact that x-rays can penetrate very thick resist on nearly arbitrary topography without scattering [1]. Typically, the resist thicknesses exposed in our laboratory are 300 nm, but this is not a fundamental limit. The ability to expose

large-aspect-ratio resist structures can be useful if a thick film of metal is lifted off. It is not necessary to add an additional gate metalization to ensure good step coverage or contact pad thickness because x-ray lithography provides the opportunity to lift off such thick layers. Consequently, one has to worry less about static charges that could destroy thin gates or about step coverage if one wants to lead thin wires onto a mesa. This last point is a problem in making planar resonant tunneling devices by e-beam lithography [46]. As can be seen in Fig. 2-15, step coverage of 100 nm lines can easily be achieved for a mesa height of about 120 nm.

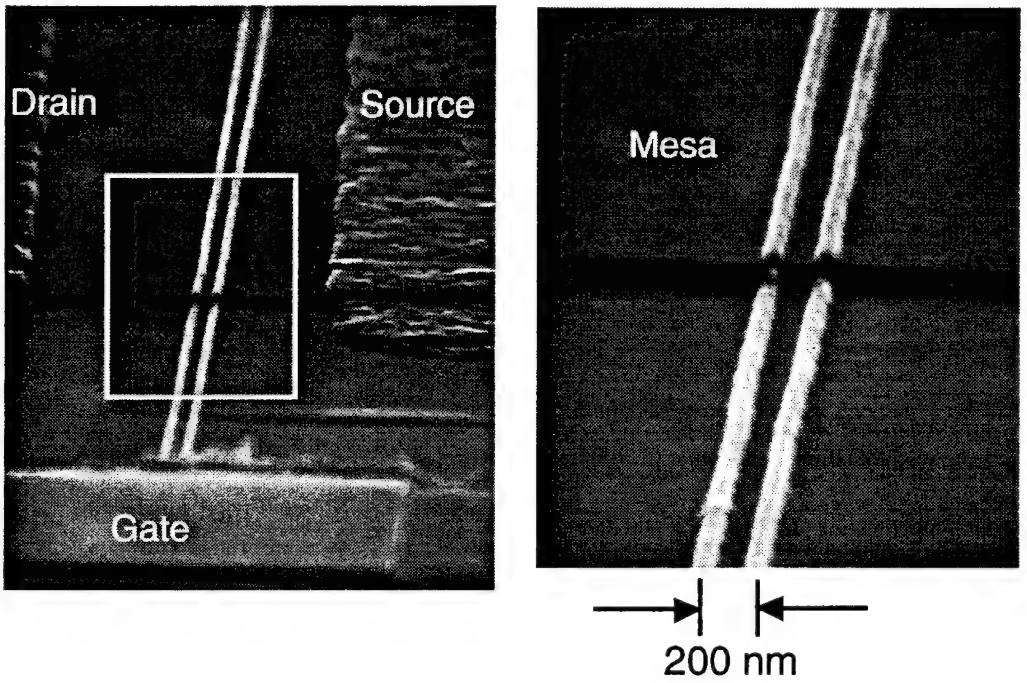


Figure 2-15: Step coverage for a two barrier planar resonant tunneling device. The gate fingers are separated by 200 nm and the metal thickness is 30 nm. The metal thickness could be made thicker. Films of more than 100 nm are frequently used in the fabrication of fine features. This improves the likelihood of good continuity across the mesa step and reduces the sensitivity of the device to static charges.

It was already mentioned that one can commit extensive time and effort into writing an x-ray mask using e-beam lithography because subsequent processing steps are parallel. This advantage has been used in the design and fabrication of optoelectronic devices, such as channel-dropping filters [16]. A "segmented fiducial grid", a unique

tool for obtaining long range coherence with an e-beam lithography tool [47], was used to write gratings onto an x-ray mask. Optoelectronic devices, produced using this technique have significant advantages concerning long range coherence compared to devices made using other techniques.

Damage during processing

In order to see quantum effects, it is crucial not only to use high quality substrates but also to maintain the high quality during processing. One of the best indicators for substrate quality is the low-temperature mobility. The higher the mobility in a particular sample, the higher the temperature at which quantum effects may be observed. It is clear that mobility degradation by exposure to high energy particles during device fabrication must be avoided.

Damage caused by electron beam evaporation of gate metals on HEMT structures [44] and on bulk materials [48–50] has been widely reported. The most important sources of radiation inside an electron beam evaporator are electrons and x-rays. Damage caused by x-ray emission during evaporation has been previously suggested as a possible mechanism in silicon systems [51]. In other experiments on GaAs [48], it was shown that through proper shielding techniques, electron irradiation could be minimized. Any damage was then attributed to stray electrons back-scattered from the evaporation target. Degradation of 2DEG mobility by e-beam lithography in the voltage range of 5-12.5 keV had previously been observed [52]. However, at higher energies (up to 20 keV), the authors did not observe any degradation.

We observed significant mobility degradation in modulation-doped AlGaAs/GaAs Hall bars after irradiation by 50 keV electrons. Unirradiated samples on the same header showed no degradation. Similarly, no degradation was observed when devices were exposed to Cu-L x-rays in a helium atmosphere [45]. These results are consistent with reports of mobility degradation due to metallization of GaAs in some e-beam evaporators.

We did not study the effect of e-beam evaporation on device quality in the fabrication of our devices. It would be difficult to replace the e-beam evaporator with a thermal evaporator for the ohmic contact evaporation because of the multilayer structure used in the ohmic contacts. In addition, the damage during this evaporation should not cause significant problems because of the annealing step that follows the evaporation. It would be of interest, however, to evaporate the gate metal with a thermal and an e-beam evaporator and compare device characteristics. Instead of low temperature mobility, one could also measure ballistic transport through quantum point contacts, which could give an indication of scattering centers introduced during processing.

Chapter 3

Device Operation and Measurements

3.1 Introduction

Reduced dimensional systems are of great importance in physics and engineering. Most of mesoscopic physics is based on such systems. The physics community has ventured into the two-, one- and zero-dimensional regimes and explored exciting new realms of physics. The engineering community, however, has not yet moved below two-dimensional systems in large-scale applications, of which the most important one is the metal oxide semiconductor field effect transistor (MOSFET). But with increasing integration density and reduced device dimensions, there is an end in sight to this route. Increased device density and clock rate on a chip already lead to problems concerning heat dissipation on a chip. The cooling of modern CPUs is one of the most important issues to be addressed in the design of a circuit board. It is likely that this problem will set a fundamental limit on both the functionality and the speed of computers. Novel devices might offer a solution to these problems, necessitating continued research in areas such as quantum-effect devices.

3.1.1 Reduced Dimensional Systems

Two dimensions

In a two-dimensional electron gas (2DEG), electrons are constrained to move in only two directions by a potential well which prevents them from moving in the third. A 2DEG can be created at a silicon-silicon dioxide interface, as is the case for the MOSFET, or at the interface between materials of different bandgap, as is the case for modulation-doped field-effect transistors (MODFETs). By far the most popular MODFET is based on the GaAs-AlGaAs material system. GaAs and AlGaAs are closely matched in their lattice constant, and one material can easily be grown on the other without creating large dislocation densities. The MODFET structure is the basis for the devices that were used in the measurements.

The electron density in the two dimensional electron gas can be varied by applying a voltage to a gate located on top of the layer structure. The energy E of the conduction electrons as a function of the two dimensional wave vector variable $\mathbf{q}$ in a 2DEG is

$$E(\mathbf{q}) = \frac{\hbar^2 \mathbf{q}^2}{2m^*} \ , \tag{3.1}$$

where m^* is the effective mass of the electron which is smaller than the electron mass, m_e, due to interaction with the lattice. For GaAs, we have $m^* \sim 0.067 m_e$. The wave vector variable, $\mathbf{q}$, lies in the x-y plane if z is the coordinate of confinement, and its magnitude is $|\mathbf{q}| = \sqrt{q_x^2 + q_y^2}$. In such a system, the two-dimensional density of states, n, in the absence of a magnetic field is

$$n_{2D}(E) = \frac{m^*}{\pi \hbar^2} \ . \tag{3.2}$$

This means that the density of states is independent of energy up to the energy of the first excited state in the z-direction. The Fermi energy of a 2DEG can be calculated

easily if the sheet electron density n_S is known as

$$E_F = \frac{n_S}{n_{2D}} = \frac{\hbar^2 k_F^2}{2m^*} \ . \tag{3.3}$$

The Fermi wavevector, k_F, is then

$$k_F = \sqrt{2\pi n_S} \ . \tag{3.4}$$

At low temperatures, phonon scattering is negligible, and electrons scatter only due to ionized impurities. An important figure of merit for a given layer structure is the elastic scattering length, the distance that an electron can travel on average without scattering elastically with an impurity. To estimate the elastic scattering length, l_e, of electrons in the 2DEG one uses the Drude formula (with the lifetime $\tau = l_e/v_F$) and relates the elastic scattering length to the Fermi wavevector and the mobility of the substrate, μ,

$$l_e = \frac{\mu\hbar k_F}{e} = \frac{\hbar\mu}{e}\sqrt{2\pi n_S} \ . \tag{3.5}$$

For the layer structure depicted in Fig. 2-13, a mobility of about 250,000 $\mathrm{cm}^2/\mathrm{Vs}$ and an electron density of about $5 \times 10^{11}\,\mathrm{cm}^{-2}$ results in an elastic scattering length of about 3μm and a Fermi wavelength of about 35 nm. Layer structures with mobilities above $10^7\,\mathrm{cm}^2/\mathrm{Vs}$ have been reported [53], leading to elastic scattering lengths in the tens of microns. The structure used in the experiments reported here represents a trade-off between high mobility and control over the 2DEG due to its close proximity to the surface.

One dimension

In a one-dimensional conductor, an additional potential barrier confines the electrons to move in only one direction. For example, if the movement in the x-direction is

confined by a potential $V(x)$, the energy is given as

$$E_j(q_y) = E_j(k_x) + \frac{\hbar^2 q_y^2}{2m^*} \ . \tag{3.6}$$

The wavevector $q_x = k_x$ is quantized and only q_y remains as a continuous variable. If the potential well is parabolic, which is a reasonably good approximation ($V(x) = m^*\omega_0^2 x^2/2$), the energy levels are equidistant with $E_j = (j - \frac{1}{2})\hbar\omega_0$. The index j is called the subband index and indicates how many conducting channels are open in the confined region. If only states in the band $j = 1$ are occupied, one refers to one-dimensional conduction, and for $j > 1$ to quasi-one dimensional conduction. One-dimensional conduction in devices has been dealt with extensively in the literature [54–57]. For a review, see refs. [58,59].

A one-dimensional conductor can be fabricated using a 2DEG by confining it with two metal gates on top of the layer structure, leaving a narrow gap of less than 0.5 μm between them. Such a device is called a quantum point contact or split gate device. By applying a negative voltage to the gates, the 2DEG is depleted underneath and only a narrow channel between the gates is left open. The optimum distance between the gates to observe subband quantization depends on the layer structure and depth of the 2DEG below the surface [60]. The length over which one-dimensional conduction can be observed depends on the elastic mean free path of the electron, which in turn depends on the layer structure and processing steps. To observe ballistic transport in a quantum point contact, the length of the one-dimensional region has to be small compared to the elastic mean free path.

The conductance of a ballistic wire at T=0 can be calculated in the following way: a voltage drop V, applied across the constriction, causes a difference in the chemical potential μ across the quantum point contact,

$$eV = \mu_1 - \mu_2 = \Delta\mu \ . \tag{3.7}$$

This difference in chemical potential leads to a current due to the uncompensated electron states on one side of the quantum point contact. The current I can be calculated as

$$I = e \sum_{j,k=1}^{N} \frac{1}{2} n_{1D,j}(E) v_j(E) T_{j,k}(E) \Delta\mu \ . \tag{3.8}$$

Here, N is the number of occupied subbands, $n_{1D,j}(E)$ is the one-dimensional density of states in the jth subband, $v_j(E)$ is the electron velocity, and $T_{j,k}(E)$ is the transmission coefficient for an electron from the kth into the jth subband. This formula can be simplified dramatically since the one-dimensional density of states $n_{1D,j} = 2(dE_j/dq)^{-1}/\pi$ is inversely proportional to the velocity of the electron $v_j(E) = 1/\hbar \, dE_j/dq$. For small applied voltages, the conductance G reduces to

$$G = \frac{2e^2}{h} \sum_{j,k=1}^{N} T_{jk} \ . \tag{3.9}$$

This is the Landauer Formula for the two terminal resistance of a quantum wire [61–63]. In the case of ballistic wires where no scattering takes place $T_{j,k}$ becomes a unity matrix, and the conductance is

$$G = \frac{2e^2}{h} N \ . \tag{3.10}$$

The conductance in a ballistic wire depends only on the number of occupied subbands and not on the geometry, *e.g.* length, as long as the wire remains ballistic. The conductance is quantized and the step height is

$$\frac{2e^2}{h} = 77.48\mu S = (12906\Omega)^{-1} \ . \tag{3.11}$$

If one biases the quantum point contact such that the chemical potential lies below the first subband, no conduction is possible, except for tunneling. If two of these barriers are biased to produce an island of electrons separated from the leads, a zero-dimensional electron gas results, as will be discussed in the following section. This arrangement is also frequently called a quantum dot, Coulomb blockade device, or

single electron transistor.

Zero dimensions

In 1988, John Scott-Thomas, then a graduate student at MIT, fabricated devices featuring long, one-dimensional conductors on silicon using x-ray lithography [64], depicted in Fig. 3-1. These devices were not expected to exhibit ballistic conduction

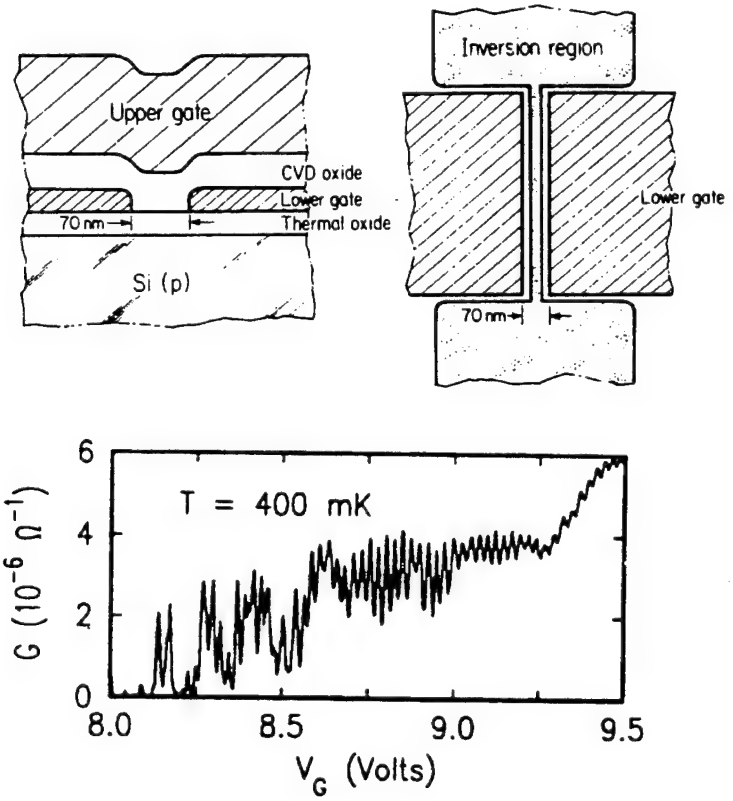


Figure 3-1: Devices in which the discovery of Coulomb blockade in semiconductor devices was made. The bottom gate was used to define a long, diffusive wire in the silicon, while the top gate was used to sweep the chemical potential within the wire. The graph on the bottom shows the conductance of the wire. After ref. [65].

characteristics, since the mobility in the silicon was not very high. He was studying the transport properties of long wires in the diffusive regime, where the elastic scattering length is longer than the width but much shorter than the length of the wire.

He measured the conductance of these wires in a dilution refrigerator and expected

to see evidence of conductance peaks associated with the population of subbands in the wire. In some of the wires, he instead observed conductance oscillations which were periodic as a function of the top gate voltage, which was used to sweep the chemical potential within the wire [65]. According to our current understanding, in each of the wires that showed these oscillations there were two impurities that produced tunneling barriers within the wires by pinching off the first subband [66]. In effect, an electron island was produced and Scott-Thomas observed the filling of the island with electrons, with each conductance peak indicating the addition of one electron to the island. This effect had been predicted and explained by Glazman and Shekhter [67], who applied knowledge of the Coulomb blockade in superconducting tunneling junctions [68] to zero-dimensional semiconductor structures.

An intuitive picture for the explanation of conductance peaks as a function of gate voltage is presented in Fig. 3-2. The schematic shows a conducting, quasi-one-dimensional channel that is interrupted twice, isolating a dot which can be accessed through tunnel barriers to two leads. The conductance of each tunneling barrier has to be adjusted below the conductance quantum of $2e^2/h$. If both tunneling barriers are biased simultaneously at these values, periodic conductance oscillations can be observed by sweeping the chemical potential within the dot using the gate. The point of Fig. 3-2 is that the charge in the dot, rather than being a continuum, is quantized to an integer number of electrons. Thus one has a staircase with step length e/C_g, rather than the classical line with slope C_g in the Q-V plot of Fig. 3-2. At each step, electrons can enter and leave the dot, resulting in current peaks.

Using artificial structures that were designed to have adjustable tunneling barriers, the interpretation of the observed conductance oscillations as single electron tunneling, proposed by van Houten [66] and Glazman [70], was soon confirmed [71–73]. One of the first such measurement, made by U. Meirav, is shown in Fig. 3-3. An inverted heterostructure was used to sweep the chemical potential in quantum dots of various sizes. The conductance oscillations as a function of gate voltage depend on the spatial

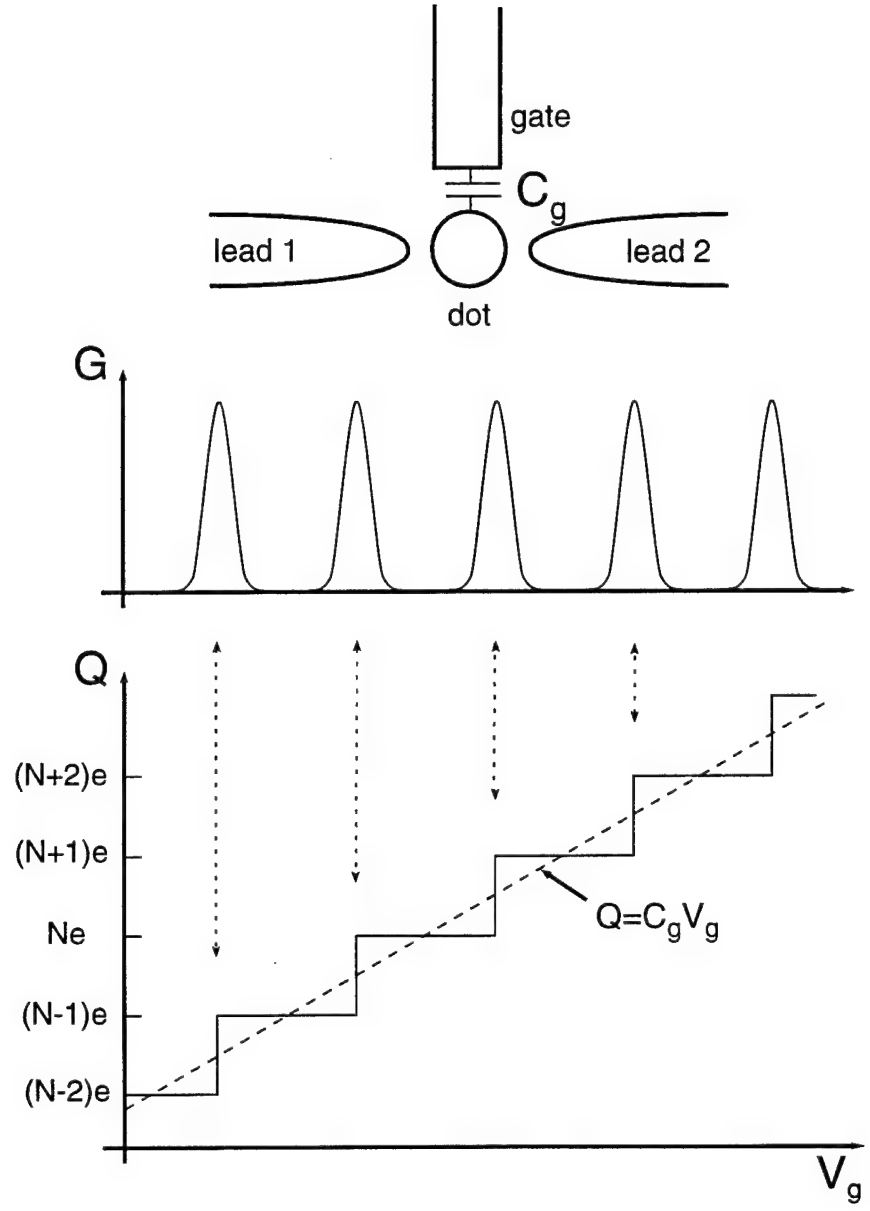


Figure 3-2: Explanation of Coulomb oscillations using the total charge of the dot. After ref. [69].

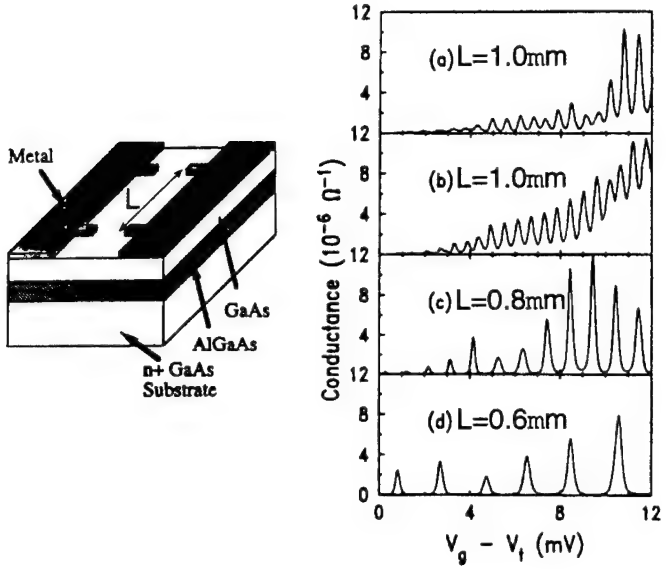


Figure 3-3: Schematic of device and conductance curves for quantum dots of different spatial extent. The device was made on an inverted heterostructure, allowing modulation of the chemical potential from the backside and control of the tunneling barriers using the metal gates. The period of conductance oscillations depends on the distance between tunneling barriers. After ref. [71].

extent of the quantum dot. There are several interesting review articles [74–77] and theses [69,73,78] on this subject.

To observe the Coulomb blockade effect and the associated conductance oscillations, the thermal energy should be small compared to the charging energy,

$$k_B T \ll \frac{e^2}{2C_\Sigma} \ , \tag{3.12}$$

where C_Σ is the sum of all capacitances to the dot, k_B is Boltzmann's constant and T the absolute temperature. Due to the capacitances involved in semiconductor structures and the limitations imposed by eq. (3.12), most measurements of this Coulomb blockade effect are done at cryogenic temperatures, often in the mK-range. There is, however, an interest to apply these new effects in engineering. The phenomenon that only an integer number of electrons can tunnel into the quantum dot has been used in so-called turnstile devices [79] to produce quantized current plateaus $I = ef$, where

f is the frequency by which the tunneling barriers are modulated. This device could be used as a current standard. Other authors have proposed logic families [68,80] and storage circuits [81] using the Coulomb blockade effect. Practical applications have not been implemented yet since low temperatures are required to observe the desired characteristics. Researchers are trying to increase the temperatures at which Coulomb blockade can be observed by reducing capacitances of the quantum dot to its environment [81,82]. One of the most exciting recent developments is the apparent discovery of Coulomb blockade in short-channel silicon on insulator (SOI) MOSFETs [83]. The authors used the pattern-dependent oxidation of silicon to produce the tunneling barriers. One of the advantages of SOI technology is the reduction of capacitances, which also happens to be the requirement for the operation of a high temperature quantum dot. In these devices, the total capacitance of the dot was about 2 aF, and strong Coulomb blockade oscillations as a function of gate voltage were observed at 77 K and conductance modulations remained *even at room temperature.*

This chapter continues with a derivation of the conductance condition for the single quantum dot in Sec. 3.2. The derivation is kept general so it can be used for comparison during the derivation for the double dot in Sec. 3.5, which is done in close analogy. A short description of the measurement setup and the probe used for low temperature measurements are given in Sec. 3.3. Section 3.4 introduces the device that was measured in the experiments and discusses the results of the single quantum dot measurements. After discussing the physics of the double dot in Sec. 3.5, measurements are shown and discussed in Sec. 3.6.

3.2 The Quantum Dot

Conduction through the quantum dot can occur only if the probability of having N electrons in the dot is just as high as having $N-1$ electrons in the dot. The probability

of having N electrons in the dot is given by the Grand Canonical Ensemble [76,84]

$$P(N) = \exp(-\frac{F(N) - N\mu}{k_B T})/\Xi \ , \tag{3.13}$$

where Ξ is the grand partition function (a constant in this case), μ is the chemical potential, F is the free energy, k_B is the Boltzmann constant and T is the temperature. At very low temperature, the probability $P(N)$ is negligible for all but one value of N, the one that minimizes $(F(N) - N\mu)$. Conductance is possible only if $P(N)$ and $P(N-1)$ are both nonzero for some N. This is the necessary condition for tunneling to take place. If this is the case, a current flows through the dot, and the number of electrons alternates between N and $N-1$. This implies that

$$F(N) - F(N-1) = \mu(N) \ . \tag{3.14}$$

At zero temperature, the free energy equals the ground-state energy of the dot, which is the sum of the charging energy U and a contribution due to the energy level quantization within the dot,

$$F(N) = U(N) + \sum_{p=1}^{N} E_p \ . \tag{3.15}$$

For the moment, we neglect the energy level quantization in the dot, which is not a valid assumption if we are dealing with extremely small quantum dots of lateral dimensions of the order of the Fermi wavelength. This case was treated in detail by Beenakker [85] and will be left aside for the moment. Its implications will be discussed in Sec. 3.2.1. Using eqs. (3.14) and (3.15), the chemical potential is defined as the energy cost of adding the Nth electron to a dot with $(N-1)$ electrons,

$$\mu(N) = U(N) - U(N-1) \ . \tag{3.16}$$

For a change in charging energy δU as a function of charge in the dot Q_d and the electrodes Q_i, we can write

$$\delta U = V_d \delta Q_d + \sum_i V_i \delta Q_i \ , \tag{3.17}$$

where V_i are the potentials of the electrodes, and V_d is the potential in the dot. Unfortunately, we have control only over the voltages V_i and not over the charges Q_i. It is desirable to interchange the dependence of the quantities in the conjugate pair. To accomplish this, we use the Legendre transformation to define a co-charging energy $\tilde{U}$ so that $\tilde{U} = U - \sum_i V_i Q_i$ and we get

$$\delta \tilde{U} = V_d \delta Q_d - \sum_i Q_i \delta V_i \ .$$

It is a straightforward exercise to transform $\tilde{U}$ to U since all the essential information is preserved during the Legendre transformation. For now we will rename $\tilde{U}$ as the charging energy U. To calculate the magnitude of the charging energy, we integrate over all charges in the dot with all electrode voltages constant ($\delta V_i = 0$),

$$U(Q_d) = \int_0^{Q_d} V(Q'_d, \{V_i\}) \delta Q'_d \ . \tag{3.18}$$

The total charge in the dot, Q_d, is related to the potentials on the electrodes V_i by the capacitance matrix elements c_{di}

$$Q_d = \sum_i c_{di} V_i \ . \tag{3.19}$$

The capacitance matrix elements c_{di} are called self capacitances for $d = i$ and mutual capacitances for $d \neq i$. The connection to circuit parameter capacitances will be

treated later on. The potential within the dot is then given as

$$V_d = \frac{1}{c_{dd}}[Q_d - \sum_{i \neq d} c_{di} V_i] = \frac{Q_d}{c_{dd}} - \sum_i \frac{c_{di}}{c_{dd}} V_i \ . \tag{3.20}$$

The self capacitance c_{dd}, according to electrostatic theory, turns out to be identical to the sum off all capacitances to the dot, introduced as C_Σ in eq. 3.12. This will be discussed later. The total charging energy can be calculated by adding up contributions of all charges in the dot using eq. (3.18),

$$U(Q_d) = \int_0^{Q_d} [\frac{Q'_d}{c_{dd}} - \sum_i \frac{c_{di}}{c_{dd}} V_i] dQ'_d = \frac{Q_d^2}{2c_{dd}} - Q_d \sum_i \frac{c_{di}}{c_{dd}} V_i \ . \tag{3.21}$$

If we sweep only one electrode potential, V_g, the charging energy U increases linearly with this potential. One can plot the energy, as is done in Fig. 3-4, for several values of N and recognize that the charging energy of the dot can be determined easily under the condition that the charging energy is minimized for each value of gate voltage. We can pick out the lines that represent $N = 3$ and $N = 4$ electrons, as shown in

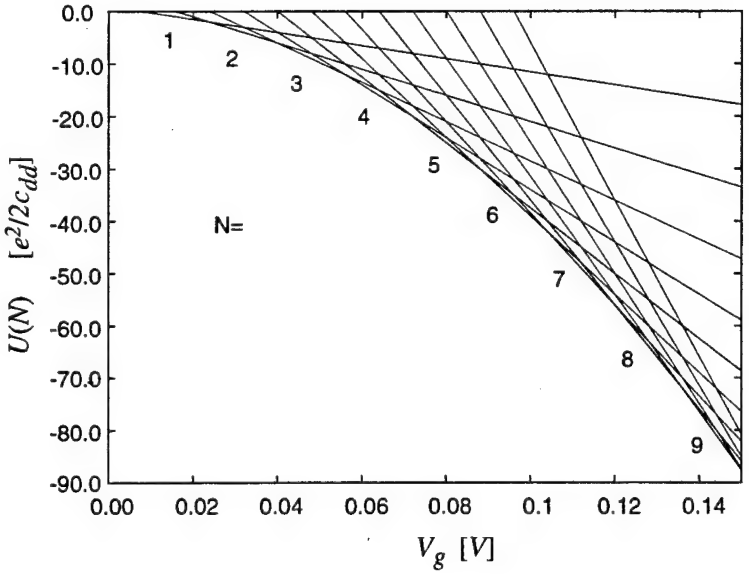


Figure 3-4: Charging Energy of the dot as a function of gate potential for $C_{dg} = 10\,\mathrm{aF}$. The system is in the state that minimizes the charging energy. After ref. [75].

Fig. 3-5, and see that at the point of intersection it is energetically just as favorable to have 3 or 4 electrons in the dot. At this value of the gate potential, we expect to

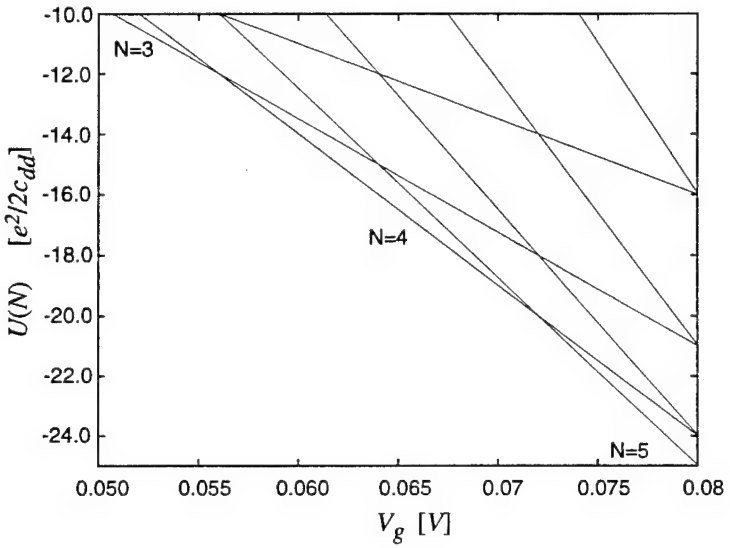


Figure 3-5: Close-up of Fig. 3-4 for $N = 4$ electrons.

see a conductance peak, since the dot can oscillate between 3 and 4 electrons, and current can flow. If the gate potential is between points of intersection, the current is suppressed by the Coulomb blockade.

Using eqs. (3.16) and (3.21), we can calculate the chemical potential as

$$\begin{aligned} \mu(N) &= U(-eN) - U(-e(N-1)) \\ &= \frac{e^2}{c_{dd}}(N - \frac{1}{2}) + \sum_i eV_i \frac{c_{di}}{c_{dd}} \end{aligned} \tag{3.22}$$

This is the chemical potential as a function of electrode voltages for a single dot. The chemical potential is swept as a function of gate voltage V_g by a factor ec_{dg}/c_{dd}. The ratio of gate capacitance to self capacitance is called the capacitive lever-arm. It is a measure of the amount by which the chemical potential in the dot is swept for a change of the potential of a gate.

To calculate the linear-response conductance G of the quantum dot, we use $G = I/V$ for the limit of $V \to 0$. In this case, source and drain are at zero potential, and

$\mu = 0$ is the condition for current flow. A typical measurement of the quantum dot is made by sweeping only one of the electrodes, $i = g$, with all other electrodes kept constant. In this case, the gate voltages at which conductance peaks occur are

$$-\frac{V_g}{e} = \frac{N - \frac{1}{2}}{c_{dg}} + \sum_{i \neq g} \frac{V_i}{e} \frac{c_{di}}{c_{dg}} \quad . \tag{3.23}$$

This means that the conductance maxima, defined as an increment of N by one, are spaced by

$$\Delta V_g = \frac{e}{c_{dg}} \tag{3.24}$$

So far, we have only dealt with capacitances. To make the result more intuitive, we transform the capacitances into capacitors.

Circuit Parameter Representation: Let each electrode j be coupled to a charge Q_i by a capacitor C_{ij}, and let all charges be imaged within the system. The charge Q_i on conductor i is known when all the voltages in the system V_j and capacitance matrix elements c_{ij} are known [86]:

$$Q_i = \sum_j c_{ij} V_j \tag{3.25}$$

Now consider the charge on the conductor i when all voltages except the voltage on conductor j are zero. The charge on conductor i is then:

$$Q_i = c_{ij} V_j \tag{3.26}$$

A positive voltage on conductor j induces a negative charge on conductor i. This means that for the mutual capacitances, $c_{ij} < 0$. However, in the case of self-capacitances, $i = j$, we have $c_{ii} > 0$ since a positive voltage induces a positive charge on the same conductor. Electrostatic theory leads us to the following relationships between the capacitance matrix elements and the circuit parameter capacitances, which

are defined to be positive numbers:

$$\begin{aligned} c_{di} &= -C_{di} \quad \text{for } i \neq d \\ c_{dd} &= +\sum_i C_{di} = C_\Sigma \quad . \end{aligned} \tag{3.27}$$

Using eq. (3.22), the chemical potential as a function of electrode voltages $\{V_i\}$ is

$$\mu(N) = \frac{e^2}{C_\Sigma}(N - \frac{1}{2}) - \sum_i eV_i \frac{C_{di}}{C_\Sigma} \quad , \tag{3.28}$$

and, using eq. (3.23), the gate voltages at which the conductance peaks occur are

$$\frac{V_g}{e} = \frac{N - \frac{1}{2}}{C_{dg}} - \sum_{i \neq g} \frac{V_i}{e} \frac{C_{di}}{C_{dg}} \quad . \tag{3.29}$$

The gate voltage difference between adjacent conductance peaks is, similar to eq. (3.24),

$$\Delta V_g = \frac{e}{C_{dg}} \quad . \tag{3.30}$$

3.2.1 Energy Level Separation within the Dot

In a more rigorous treatment, one has to take the energy levels within the quantum dot into account, which arise due to the small spatial extent of the dot [85]. Using eq. (3.15), the conductance condition becomes

$$U(N) - U(N-1) = \mu(N) - E_N \quad . \tag{3.31}$$

The chemical potential in the dot is then

$$\mu(N) = E_N + U(N) - U(N-1) = E_N + (N - \frac{1}{2})\frac{e^2}{C_\Sigma} - \sum_i eV_i \frac{C_{di}}{C_\Sigma} \quad , \tag{3.32}$$

and the gate voltage for a conduction peak is

$$\frac{V_g}{e} = \frac{N - \frac{1}{2}}{C_{dg}} - \sum_{i \neq g} \frac{V_i}{e} \frac{C_{di}}{C_{dg}} + \frac{C_\Sigma}{C_{dg}} \frac{E_N}{e^2} \quad . \tag{3.33}$$

The distance between adjacent conductance peaks is

$$\Delta V_g = \frac{e}{C_{dg}} + \frac{C_\Sigma}{C_{dg}} \frac{\Delta E}{e} \quad . \tag{3.34}$$

The spin degeneracy of the levels is lifted by the charging energy. The distance between adjacent peaks alternates between e^2/C_{dg} and $e^2/C_{dg} + (C_\Sigma/C_{dg})(\Delta E/e)$.

If the energy difference ΔE is large compared to the thermal energy, only one energy level within the well contributes to conduction and one approaches the limit of resonant tunneling through a well. If ΔE is larger than the incremental charging energy, resonant tunneling is the predominant effect. Resonant tunneling occurs if the Fermi level in the leads matches the energy level of a bound state within the well.

3.2.2 Thermal Spread

The behavior of the conductance peaks with temperature was investigated by Beenakker [85]. Two limiting cases can be given for the conductance as a function of chemical potential, depending on the size of the energy level separation ΔE compared to the thermal energy $k_B T$. First, we consider the lineshape in the resonant tunneling regime, $k_B T \ll \Delta E \ll e^2/C_\Sigma$. The thermal energy is smaller than the energy level separation, as can be the case for very low temperatures and small quantum dots. The conductance is

$$G = -e^2 \frac{\Gamma_p^l \Gamma_p^r}{\Gamma_p^l + \Gamma_p^r} f'(\mu_p) = \frac{e^2}{4 k_B T} \frac{\Gamma_p^l \Gamma_p^r}{\Gamma_p^l + \Gamma_p^r} \cosh^{-2}\left(\frac{\mu_p}{2 k_B T}\right) \quad , \tag{3.35}$$

where f' is the derivative of the Fermi-Dirac function. Γ^l and Γ^r are the tunneling rates for the left and right barriers, respectively. For a sweep of gate voltage, we fit

the conductance using the formula

$$G(V_g) = \sum_i G_{ri} \cosh^{-2}\left(\frac{e\alpha(V_g - V_{ri})}{2k_BT}\right) , \tag{3.36}$$

where V_{ri} are the position of conductance resonances with peak conductance G_{ri}. Note that the capacitive lever-arm $\alpha = C_{dg}/C_\Sigma$ causes a spreading of the peak, causing distinguishable peaks to disappear for large C_Σ.

For the classical limit, when $\Delta E \ll k_BT \ll e^2/C_\Sigma$, a continuum of energy levels contributes to the conductance and the lineshape is given by [85,87]

$$\begin{aligned} G &= G_{max}\frac{\mu/k_BT}{\sinh(\mu/k_BT)} \simeq G_{max}\cosh^{-2}\left(\frac{\mu}{2.5k_BT}\right) \\ G_{max} &= \frac{e^2}{2\Delta E}\frac{\Gamma^l\Gamma^r}{\Gamma^l+\Gamma^r} . \end{aligned} \tag{3.37}$$

For a sweep of the gate voltage, the conductance can be fit using

$$G(V_g) = \sum_i G_{ri} \cosh^{-2}\left(\frac{e\alpha(V_g - V_{ri})}{2.5k_BT}\right) . \tag{3.38}$$

Although the line shapes in the classical and resonant tunneling regimes are different, they are not distinguishable if temperature is used as a fitting parameter.

3.3 Low Temperature Measurements

A schematic of the refrigerator used to measure devices at cryogenic temperatures is shown in Fig. 3-6. Using this probe, it is possible to cool a sample to 300 mK within a few hours and to keep it at base temperature for more than 8 hours. Following is a short explanation of how the probe works. More extensive operating instructions for this particular refrigerator are given in ref. [78].

The probe is immersed in liquid ^{4}He, keeping the outside of the inner vacuum can

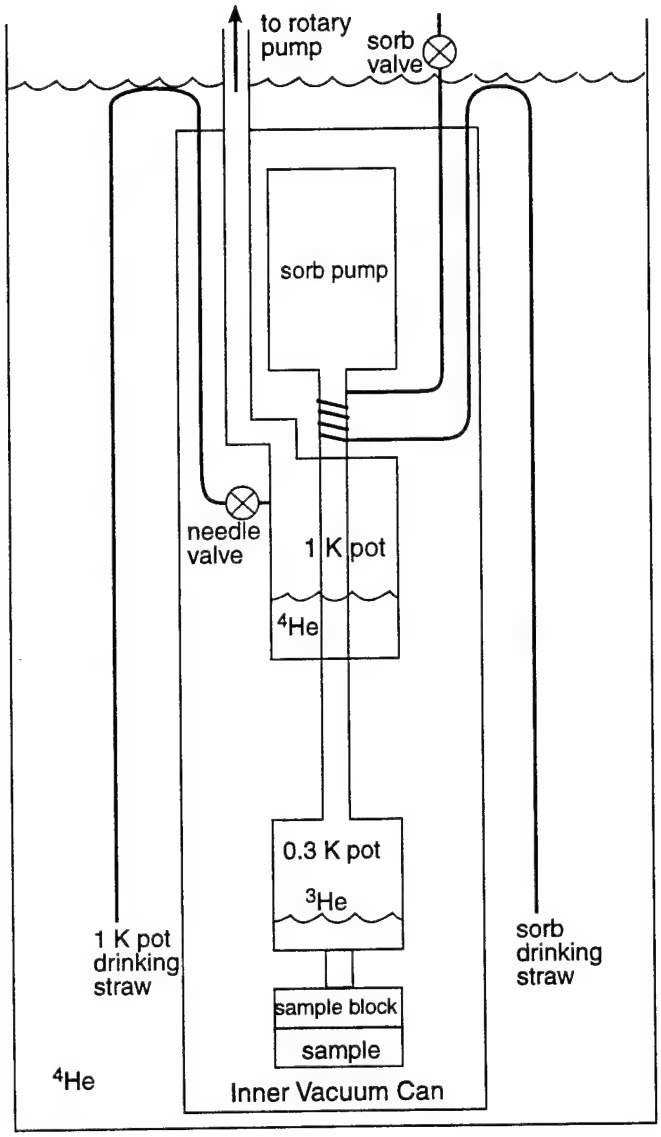


Figure 3-6: Schematic of the Oxford Instruments Heliox insertion probe by Oxford Instruments. Temperatures of 300 mK are routinely achieved with a hold time of more than 8 hours.

at a constant temperature of 4.2 K. The drinking straws are immersed, as indicated in Fig. 3-6. Cooling is achieved by reducing the vapor pressure of liquid ^{3}He using a sorb pump which adsorbs gas if cooled below a critical temperature of about 12 K. To initiate a cooling cycle, liquid ^{4}He is admitted into the 1 K pot and pumped on using an external rotary pump. This reduces the temperature of the ^{4}He to about 1.2 K. The sorb pump is then heated above its critical temperature, while the 1 K pot and the tube that connects sorb and 0.3 K pot is kept cold using the sorb drinking straw. The ^{3}He that is released during heating of the sorb is cooled to 1.2 K while passing through the 1 K pot, and condenses in the 0.3 K pot. The sorb heater is then turned off and the temperature of the 0.3 K pot, which is thermally shorted to the sample block, reaches the base temperature of about 300 mK within a few minutes. The sample is glued to a 44-pin chip carrier using silver paint and bonded with aluminum wire. The back of the chip carrier is thermally sunk to the sample block. The probe is wired with manganin leads which are thermally shorted to the 0.3 K pot. Manganin is a material of low thermal but high electrical conductivity.

The conductance of the device is measured by applying a DC or low frequency AC voltage to an ohmic contact (drain) and measuring the current through another ohmic contact (source). Because the currents are very small, it is important to minimize the noise level. This is also important because any noise on the wires can heat the sample and deteriorate device performance. For this reason, all electrical lines that are used for device and temperature measurements must be filtered. In order to measure the high-impedance devices, electronics had to be designed and built. The design of the filter box is described in Appendix H. For AC measurements, a PAR 5210 model lock-in amplifier was used. Constant gate voltages were set using battery-operated voltages sources, and the variable gate voltage was set using an HP 3325B function generator. A computer-aided data acquisition system which controls the lock-in by an IEEE-488 bus was used for the measurements.

3.4 Measurement of Quantum Dot

A device was built, consisting of four consecutive quantum point contacts which can be controlled independently. An SEM photograph of a monitor sample with identical geometry is shown in Fig. 3-7. A quantum dot is produced by biasing two quantum

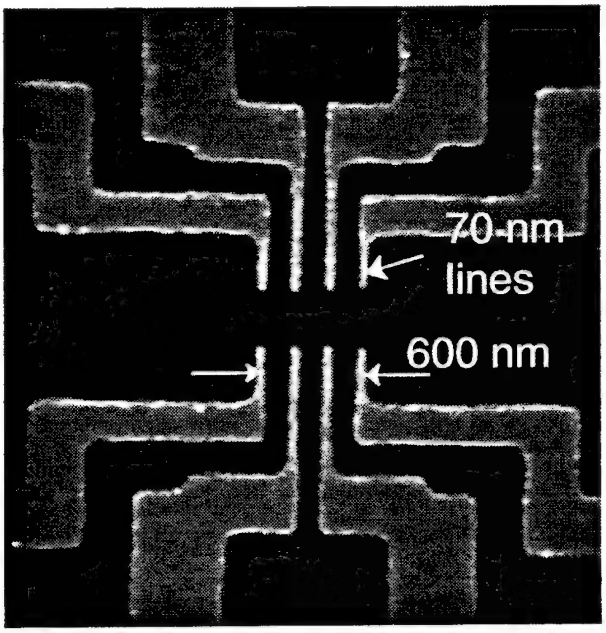


Figure 3-7: SEM photograph of the completed device. The fingers of the quantum point contacts have a width of approximately 70 nm. A quantum dot can be produced by biasing two quantum point contacts below the the first subband, thus creating a tunnel barrier for electrons in the dot.

point contacts below the first subband. This allows electron transport into and out of the dot only via tunneling. The remaining gates can be used to sweep the chemical potential within the dot. The size of the quantum dot can be adjusted from less than 200 nm to approximately 600 nm by pinching off the different pairs of quantum point contacts. A schematic showing how the different dot sizes are achieved is shown in Fig. 3-8. The finished device has minimum linewidths below 70 nm. The small linewidth causes the capacitances of the gates to the dot to be small. According to eq. (3.12), the temperatures at which observation of conductance oscillations is possible should be higher than for coarser gate structures.

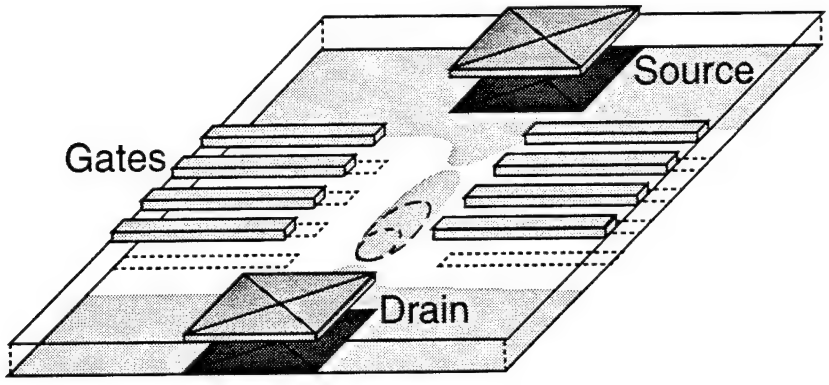


Figure 3-8: Schematic of biasing conditions for different dot sizes. Dots of various sizes can be made by pinching off two pairs of the four consecutive quantum point contacts. Each pair of quantum point contacts has an opening of 300 nm, the distance from one pair to the next is 300 nm. The maximum possible length of the dot is 600 nm.

A schematic of a typical measurement setup is shown in Fig. 3-9. An AC voltage, in this case 100 μV, is imposed on the source of the device, and the current is measured at the drain. To bias one quantum dot, two of the four pairs of quantum point contacts are biased in the tunneling regime. The remaining four gates can be used either to sweep the chemical potential in the dot or kept constant. Since all gates can be controlled independently, a large variety of experiments can be conducted. In addition to several sizes of quantum dot, as described above, coupling of small quantum dots and transitions from large to small dots can be investigated.

The conductance oscillations are strongest when both barriers are tuned to be identical in transmission. To fine-tune the barriers, the following method can be used. One of the barriers should be incrementally made stronger. If the threshold of the device shifts positive, the stronger barrier has been made stronger. If the threshold stays the same, which is the desired case, the weaker barrier has been made stronger, and this generally results in stronger conductance oscillations. An example of a conductance measurement versus gate voltage for the biasing conditions indicated in Fig. 3-9 is shown in Fig. 3-10.

It should be noted that due to the small spatial extent of the device it is not trivial to bias a dot. Large Coulomb blockade devices can be biased by determining

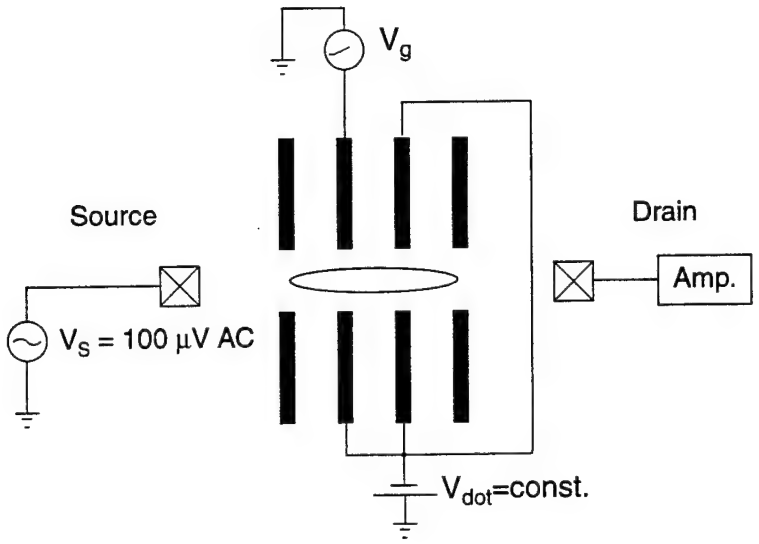


Figure 3-9: Schematic of the measurement setup used to measure conductances.

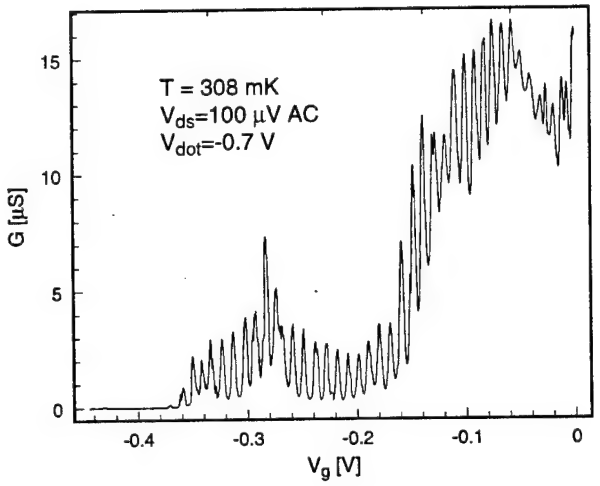


Figure 3-10: Plot of conductance versus gate voltage for a quantum dot, having a length of 600 nm. The two outermost quantum point contacts define the dot. For comparison of scales, the value of quantized conduction is $2e^2/h \simeq 77\mu$S.

the voltage of the individual quantum point contact at which only tunneling takes place, and then using this voltage to bias a quantum dot. The close proximity between the gates in this structure causes a strong influence of the change of one tunneling barrier on an adjacent barrier. For this reason, biasing has to be done by iteration for very small quantum dots. Another consequence of this coupling of tunneling barriers to nearby electrodes is that an electrode that is used to sweep the chemical potential within the dot also alters the tunneling barriers and shape and size of the dot. This means that conductance oscillations can be seen over smaller ranges of gate voltage than for larger devices.

3.4.1 Gate Capacitances

For the largest possible dot, with 600 nm distance between the center of the tunneling barriers, four gates access the middle of the dot. Each and any combination of these four gates can be swept and reveals a capacitance for the particular gates to the quantum dot. Figure 3-11 shows a plot of the conductance for the the dot with one, two, three, and four gates simultaneously swept. For two gates, adjacent gates were swept simultaneously. This experiment proves that the gates are independent of each other and that each gate finger contributes approximately the same capacitance of about 16 aF to the dot-to-gate capacitance. All four center gates, swept independently, show an almost identical periodicity of conductance oscillations, which indicates close matching of their geometries.

3.4.2 Measurement of a Variable Size Quantum Dot Device

Different-sized dots, indicated in Fig. 3-8, can be measured by biasing different quantum point contacts for each measurement. The conductances for 600, 400, and 200 nm distance between the defining quantum point contacts is plotted in Fig. 3-12 as a function of gate voltage. In case of the 200 nm dot, a gate outside the tunneling barriers was swept. The actual size of the dots is smaller than the distance between quantum

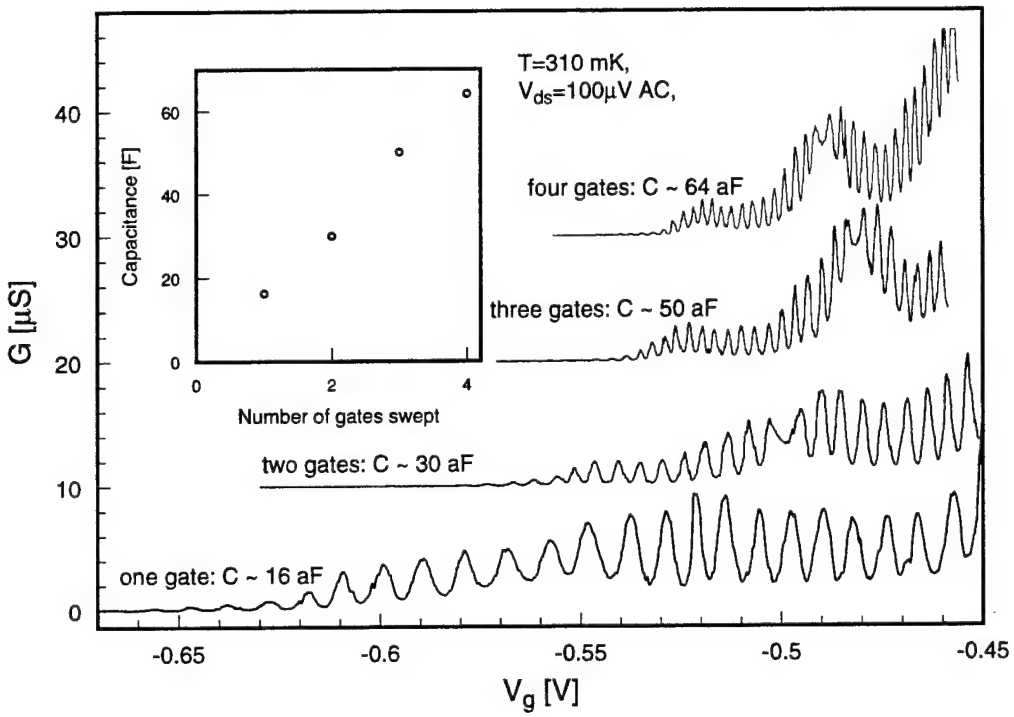


Figure 3-11: Conductance of a 600 nm dot with one, two, three and all four inner gates swept.

point contacts by approximately half the length of both tunneling barriers.

Considering that the capacitance is related to the period of conductance oscillations, as given in eq. (3.30), we can assign a dot-to-gate capacitance for the 400 nm dot of $\simeq$8 aF and for the 600 nm dot of $\simeq$16 aF. For the smallest dot, we observe that the spacing between the maxima is not constant, possibly indicating a sufficiently large energy level separation according to eq. (3.34). The capacitance is then determined by the smallest spacing between conductance peaks and can be estimated to be $\simeq$2 aF.

Fit with Fermi Function

A fit for the conductance of the 400 and 200 nm dot with the lineshape given in eq. (3.36) is shown in Fig. 3-13. It should be noted that the fit is only reasonably good for the first few conductance peaks. For the remaining peaks, the off-resonance

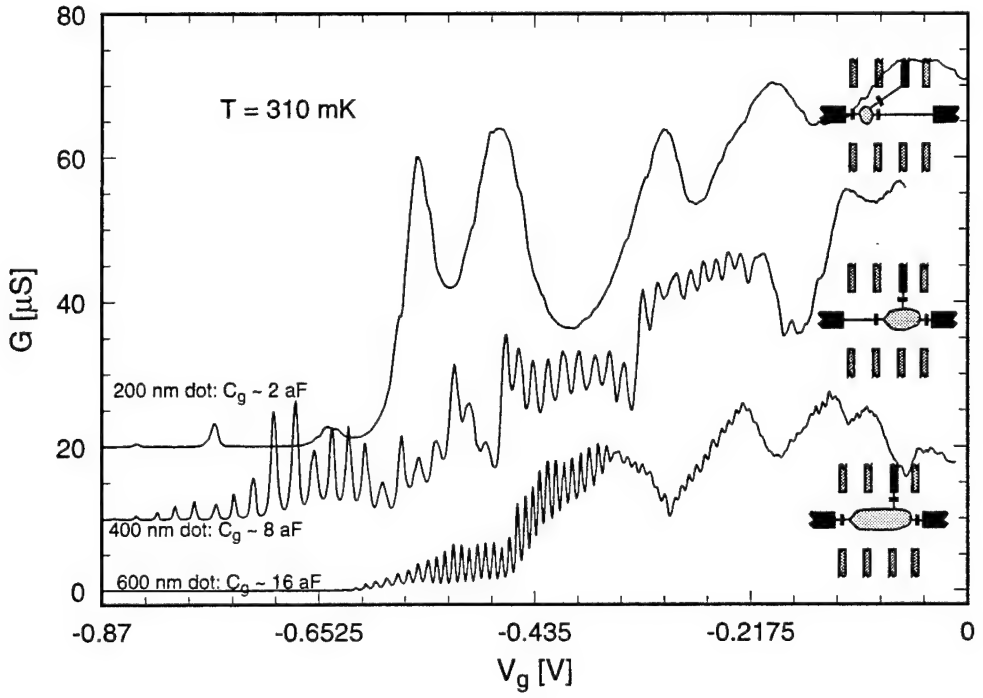


Figure 3-12: Conductance of quantum dots with separation of 600, 400, and 200 nm between tunneling barriers. In all cases only one of the gates, indicated as the black gate in the inserts, is swept. In the case of the 200 nm dot a gate beyond the tunneling barrier was swept. The curves are offset by 10 μS for clarity.

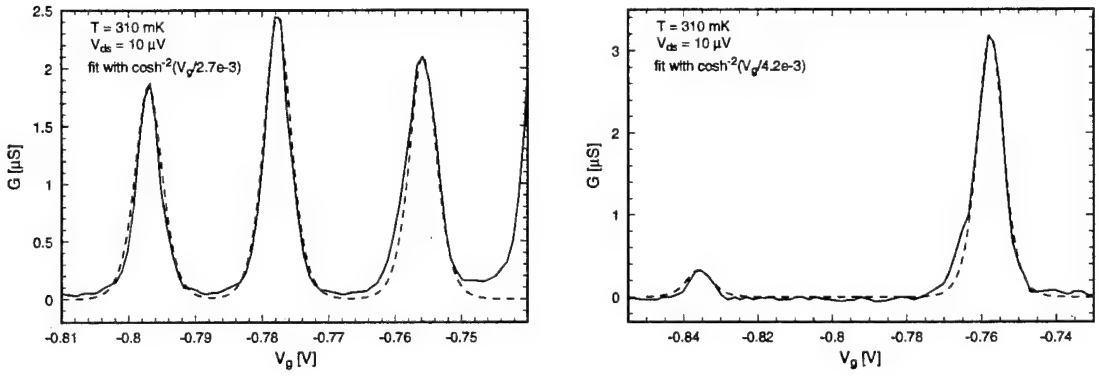


Figure 3-13: Fit of measured conductances for the 400 and 200 nm dot with eq. (3.36).

	200 nm (a)	400 nm (b)
C_{gi} [aF]	2.0	8.1
$C_{i\Sigma}[aF]$	310	400

Table 3.1: Capacitances for 200 nm dot (a) and 400 nm dot (b). The sum-capacitances are extracted from the fit and the gate capacitances from the periodicity of the conductance peaks.

conductance becomes large. The fit with the Fermi function enables us to extract an approximate value for the capacitive lever-arms of the dots. If the fit for the resonant tunneling regime is used, the capacitive lever-arm for the 200 nm dot, a, is $\alpha_a \simeq 1.3 \times 10^{-2}$, and for the 400 nm dot, b, is $\alpha_b \simeq 2.0 \times 10^{-2}$. It must be noted that the lever-arms for the system are smaller than expected. If the remaining gates have a capacitance of the same order of magnitude, the total gate capacitance for the 400 nm dot would still be below 100 aF. Even if one considers the capacitances to the leads and to the substrate, whose conducting back-gate is separated from the 2DEG by the thickness of the wafer, it is difficult to imagine the capacitance to reach 400 aF. It is possible that the experiments were conducted at higher temperature than was measured at the sample block [88]. However, the ratio between the lever-arms, important for discussion of the two dots measured in series, should be approximately correct. The values for the capacitances for the two dots are given in table 3.1. The capacitance given for dot b is extracted from the period of the first few conductance maxima. For higher gate voltages, the dot becomes larger in size, resulting in a larger capacitance of about 9 aF and higher periodicity.

3.5 The Double Dot

The device shown in Fig. 3-7 can be biased to produce two quantum dots in series by using three pairs of quantum point contacts to produce three consecutive tunneling barriers. A schematic of the system is shown in Fig. 3-14. The two dots were already measured individually in Sec. 3.4 and are denoted as a for the small, and b for the

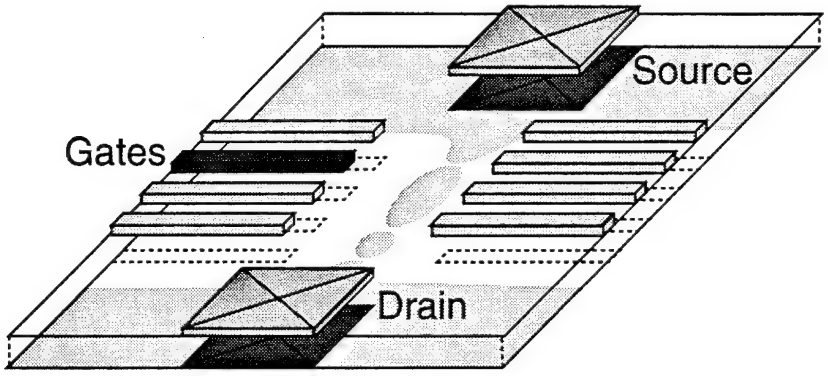


Figure 3-14: Schematic of the double dot structure. The center-to-center distance between tunneling barriers is 200 nm and 400 nm for the two dots. The gate indicated in dark is swept.

large dot.

Systems containing multiple dots have been investigated theoretically for a couple of years [89–93], but experimental studies are just emerging. Some experiments focus on identical or slightly detuned quantum dots [94], while others focus on dots with identical physical layout which are strongly detuned through gate bias [95]. The experiments described in this work use dots of unequal size and significantly unequal capacitances. If one stretches the physical analog of one quantum dot being an atom with its confined states [77], the analog of a double dot is a molecule. The coupling of the two atoms of this molecule can be modified by simply changing the gate voltage of the tunneling barrier separating the dots. We expect the physical characteristics of this double dot to be different from the characteristics of the isolated dots. It would be of great interest to map the energy level spectrum of such structures.

The derivation for the conductance condition for a double dot is done in close analogy to the derivation for the single dot of Sec. 3.2. Equations (3.17) to (3.23) for the single dot correspond to eqs. (3.39) to (3.50) for the double dot. It is hoped that this close analogy will avoid confusion over the complexity of the formulas. The meaning of the results will be further explained by considering special cases. In this derivation, we again neglect the energy level separation caused by the finite extent of

the dots, keeping in mind that this can be accounted for as in Sec. 3.2.1. This derivation considers the potentials and capacitances of both dots to an arbitrary number of electrodes. This treatment is necessary for a quantitative analysis of measurement results.

Consider a double dot, with the two dots denoted by indices a and b, surrounded by an arbitrary number of electrodes i. The change in charging energy in the system is

$$\delta U = V_a \delta Q_a + V_b \delta Q_b + \sum_i V_i \delta Q_i \ . \tag{3.39}$$

The first two terms represent a change in energy of the two dots. The charges Q_i on the electrodes are not controlled variables. To get an equation for the internal energy as a function of the controlled variables Q_a, Q_b, and V_i, we use again the Legendre Transformation $\tilde{U} = U - \sum_i V_i Q_i$ to get

$$\delta \tilde{U} = V_a \delta Q_a + V_b \delta Q_b - \sum_i Q_i \delta V_i \ . \tag{3.40}$$

We again rename $\tilde{U}$ as the charging energy U. The experiment is done with fixed $\{V_i\}$, so that

$$U|_{\{\delta V_i = 0\}} = \int_0^{Q_a} V_a(Q'_a, Q_b = 0, \{V_i\}) \delta Q'_a + \int_0^{Q_b} V_b(Q_a, Q'_b, \{V_i\}) \delta Q'_b \ . \tag{3.41}$$

To find V_a and V_b as a function of the variables Q_a and Q_b, we use again the capacitance matrix elements c_{am},

$$Q_a = \sum_m c_{am} V_m \ . \tag{3.42}$$

Solving for the potential V_a we get

$$V_a = \frac{1}{c_{aa}} [Q_a - \sum_{m \neq a} c_{am} V_m] \ , \tag{3.43}$$

and the dot potentials V_a and V_b become

$$\begin{aligned} V_a &= \frac{1}{c_{aa}}Q_a - \frac{c_{ab}}{c_{aa}}V_b - \sum_i \frac{c_{ai}}{c_{aa}}V_i \\ V_b &= \frac{1}{c_{bb}}Q_b - \frac{c_{ab}}{c_{bb}}V_a - \sum_i \frac{c_{bi}}{c_{bb}}V_i \ . \end{aligned} \tag{3.44}$$

Solving for V_a and V_b by elimination, we get

$$\begin{aligned} V_a &= \frac{1}{\kappa c_{aa}}Q_a - \frac{1}{c_{ab}}\frac{1-\kappa}{\kappa}Q_b - \sum_i \frac{V_i}{\kappa}\frac{c_{ai}}{c_{aa}}\frac{1}{\alpha_i} \\ V_b &= \frac{1}{\kappa c_{bb}}Q_b - \frac{1}{c_{ab}}\frac{1-\kappa}{\kappa}Q_a - \sum_i \frac{V_i}{\kappa}\frac{c_{bi}}{c_{bb}}\frac{1}{\beta_i} \ , \end{aligned} \tag{3.45}$$

where the parameters κ, α_i, and β_i are functions of the capacitance matrix elements and are defined as

$$\begin{aligned} \kappa &= 1 - \frac{c_{ab}^2}{c_{aa}c_{bb}} \\ \alpha_i^{-1} &= 1 - \frac{c_{ab}}{c_{bb}}\frac{c_{bi}}{c_{ai}} = 1 - (1-\kappa)\frac{c_{bi}}{c_{ai}}\frac{c_{aa}}{c_{ab}} \\ \beta_i^{-1} &= 1 - \frac{c_{ab}}{c_{aa}}\frac{c_{ai}}{c_{bi}} = 1 - (1-\kappa)\frac{c_{ai}}{c_{bi}}\frac{c_{bb}}{c_{ab}} \ . \end{aligned} \tag{3.46}$$

Using eq. (3.41), we integrate over the charge in both dots and get for the charging energy:

$$U = \frac{Q_a^2}{2c_{aa}\kappa} + \frac{Q_b^2}{2c_{bb}\kappa} - \frac{Q_aQ_b}{c_{ab}}\frac{1-\kappa}{\kappa} - \sum_i V_i(Q_a\frac{c_{ai}}{c_{aa}}\frac{1}{\alpha_i\kappa} + Q_b\frac{c_{bi}}{c_{bb}}\frac{1}{\beta_i\kappa}) \tag{3.47}$$

The chemical potential can be calculated, using eq. (3.16), as

$$\mu_a(N_a, N_b) = U(-eN_a, -eN_b) - U(-e(N_a-1), -eN_b) \ , \tag{3.48}$$

and likewise for $\mu_b(N_a, N_b)$:

$$\begin{aligned}\mu_a(N_a,N_b) &= \frac{e^2}{\kappa c_{aa}}(N_a-\frac{1}{2}) - \frac{e^2}{c_{ab}}\frac{1-\kappa}{\kappa}N_b + \sum_i eV_i\frac{c_{ai}}{c_{aa}}\frac{1}{\alpha_i\kappa} \\ \mu_b(N_a,N_b) &= \frac{e^2}{\kappa c_{bb}}(N_b-\frac{1}{2}) - \frac{e^2}{c_{ab}}\frac{1-\kappa}{\kappa}N_a + \sum_i eV_i\frac{c_{bi}}{c_{bb}}\frac{1}{\beta_i\kappa}\ .\end{aligned} \tag{3.49}$$

The conductance is measured for negligible drain voltage, so source and drain are assumed to be at zero potential. Then $\mu_a = \mu_b = 0$ give the conditions for current flow. If only one of the chemical potentials in the dots lines up with the chemical potential in the wires, tunneling into and out of the dot can still happen, but current flow is suppressed. If we define one of the electrodes i as the gate g, with voltage V_g, the conditions for current flow are

$$\begin{aligned}-\frac{V_g}{e} &= \frac{N_a-\frac{1}{2}}{c_{ag}}\alpha_g - \frac{N_b}{c_{ab}}\frac{c_{aa}}{c_{ag}}(1-\kappa)\alpha_g + \sum_{i\neq g}\frac{V_i}{e}\frac{c_{ai}}{c_{ag}}\frac{\alpha_g}{\alpha_i} \\ -\frac{V_g}{e} &= \frac{N_b-\frac{1}{2}}{c_{bg}}\beta_g - \frac{N_a}{c_{ab}}\frac{c_{bb}}{c_{bg}}(1-\kappa)\beta_g + \sum_{i\neq g}\frac{V_i}{e}\frac{c_{bi}}{c_{bg}}\frac{\beta_g}{\beta_i}\ .\end{aligned} \tag{3.50}$$

Circuit Parameter Representation: Using the formalism developed in Sec. 3.2, the mutual and self-capacitances are related to circuit elements by the relations

$$\begin{aligned}c_{ai} &= -C_{ai} \quad \text{for } i\neq a \\ c_{bi} &= -C_{bi} \quad \text{for } i\neq b \\ c_{aa} &= C_{ab}+\sum_i C_{ai} = C_{ab}+C_a \\ c_{bb} &= C_{ab}+\sum_i C_{bi} = C_{ab}+C_b\ .\end{aligned} \tag{3.51}$$

With this, the parameters κ, α_i, β_i, and a newly defined $\tilde{C}$ become

$$\begin{aligned}
\kappa &= 1 - \frac{C_{ab}^2}{(C_{ab}+C_a)(C_{ab}+C_b)} \\
\alpha_i &= \frac{(C_{ab}+C_b)C_{ai}}{(C_{ab}+C_b)C_{ai}+C_{ab}C_{bi}} \\
\beta_i &= \frac{(C_{ab}+C_a)C_{bi}}{(C_{ab}+C_a)C_{bi}+C_{ab}C_{ai}} \\
\tilde{C}^2 &= C_{ab}C_a + C_{ab}C_b + C_aC_b \ .
\end{aligned} \tag{3.52}$$

The charging energy is

$$\begin{aligned}
U = &\frac{Q_a^2(C_{ab}+C_b) + Q_b^2(C_{ab}+C_a) + 2Q_aQ_bC_{ab}}{2\tilde{C}^2} + \\
&\sum_i V_i\left(Q_a\frac{(C_{ab}+C_b)C_{ai}+C_{ab}C_{bi}}{\tilde{C}^2} + Q_b\frac{(C_{ab}+C_a)C_{bi}+C_{ab}C_{ai}}{\tilde{C}^2}\right) \ .
\end{aligned} \tag{3.53}$$

The chemical potential is then

$$\begin{aligned}
\mu_a(N_a,N_b) &= \frac{e^2(C_{ab}+C_b)}{\tilde{C}^2}(N_a-\frac{1}{2}) + \frac{e^2C_{ab}}{\tilde{C}^2}N_b - \\
&\quad \sum_i eV_i\frac{(C_{ab}+C_b)C_{ai}+C_{ab}C_{bi}}{\tilde{C}^2} \\
\mu_b(N_a,N_b) &= \frac{e^2(C_{ab}+C_a)}{\tilde{C}^2}(N_b-\frac{1}{2}) + \frac{e^2C_{ab}}{\tilde{C}^2}N_a - \\
&\quad \sum_i eV_i\frac{(C_{ab}+C_a)C_{bi}+C_{ab}C_{ai}}{\tilde{C}^2} \ ,
\end{aligned} \tag{3.54}$$

and the conditions for conduction 3.50 become

$$\begin{aligned}
\frac{V_g}{e} &= \frac{N_a-\frac{1}{2}}{C_{ag}}\frac{C_{ag}(C_{ab}+C_b)}{(C_{ab}+C_b)C_{ag}+C_{ab}C_{bg}} + \frac{N_bC_{ab}}{C_{ag}(C_{ab}+C_b)+C_{ab}C_{bg}} - \\
&\quad \sum_{i\neq g}\frac{V_i}{e}\frac{(C_{ab}+C_b)C_{ai}+C_{ab}C_{bi}}{C_{ag}(C_{ab}+C_b)+C_{ab}C_{bg}} \\
\frac{V_g}{e} &= \frac{N_b-\frac{1}{2}}{C_{bg}}\frac{C_{bg}(C_{ab}+C_a)}{(C_{ab}+C_a)C_{bg}+C_{ab}C_{ag}} + \frac{N_aC_{ab}}{C_{bg}(C_{ab}+C_a)+C_{ab}C_{ag}} - \\
&\quad \sum_{i\neq g}\frac{V_i}{e}\frac{(C_{ab}+C_a)C_{bi}+C_{ab}C_{ai}}{C_{bg}(C_{ab}+C_a)+C_{ab}C_{ag}} \ .
\end{aligned} \tag{3.55}$$

These are the general expressions for the charging energy and conductance conditions in a double dot system with an arbitrary number of electrodes. In the following paragraphs we discuss some special cases which can be deduced from these equations.

Special Case: Single Dot Limit ($C_{ab} \to \infty$) In the limit of very strong coupling we should get the solution for one single dot with $N_a + N_b$ electrons and capacitance $C_{ag}+C_{bg}$. The expressions for charging energy, chemical potential and voltage become:

$$U = \frac{(Q_a + Q_b)^2}{2(C_a + C_b)} + \sum_i V_i(Q_a + Q_b)\frac{C_{ai} + C_{bi}}{C_a + C_b} \tag{3.56}$$

$$\mu_a = \mu_b = \frac{e^2}{C_a + C_b}(N_a + N_b - \frac{1}{2}) - \sum_i eV_i\frac{C_{ai} + C_{bi}}{C_a + C_b} \tag{3.57}$$

$$\frac{V_g}{e} = \frac{N_a + N_b - \frac{1}{2}}{C_{ag} + C_{bg}} - \sum_{i \neq g}\frac{V_i}{e}\frac{C_{ai} + C_{bi}}{C_{ag} + C_{bg}} \tag{3.58}$$

It can easily be verified that the last two expressions are identical to eqs. (3.28) and (3.29), with $N_a + N_b = N$ and $C_{ai} + C_{bi} = C_i$.

Special Case: Independent Dot Limit ($C_{ab} \to 0$) In the limit of no coupling between the dots, $C_{ab} \to 0$, the expressions for the charging energy, the chemical potential and the voltage conditions for current transport become:

$$U = \frac{Q_a^2}{2C_a} + \frac{Q_b^2}{2C_b} + \sum_i V_i(Q_a\frac{C_{ai}}{C_a} + Q_b\frac{C_{bi}}{C_b}) \tag{3.59}$$

$$\begin{aligned} \mu_a &= \frac{e^2}{C_a}(N_a - \frac{1}{2}) - \sum_i eV_i\frac{C_{ai}}{C_a} \\ \mu_b &= \frac{e^2}{C_b}(N_b - \frac{1}{2}) - \sum_i eV_i\frac{C_{bi}}{C_b} \end{aligned} \tag{3.60}$$

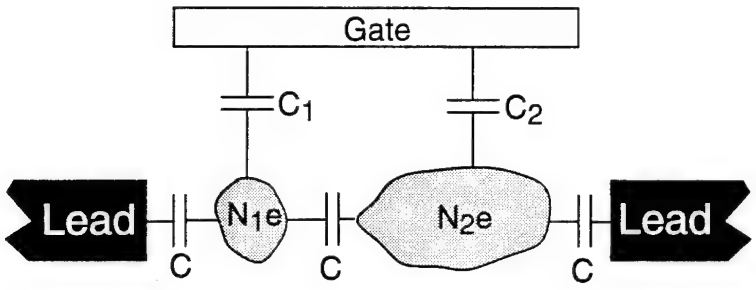


Figure 3-15: Capacitances, as defined by Ruzin.

$$\begin{aligned} \frac{V_g}{e} &= \frac{N_a - \frac{1}{2}}{C_{ag}} - \sum_{i \neq g} \frac{V_i}{e} \frac{C_{ai}}{C_{ag}} \\ \frac{V_g}{e} &= \frac{N_b - \frac{1}{2}}{C_{bg}} - \sum_{i \neq g} \frac{V_i}{e} \frac{C_{bi}}{C_{bg}} \end{aligned} \tag{3.61}$$

Note that in this case $C_a = C_{a\Sigma}$, because the coupling capacitance C_{ab} is neglected. There are two conditions that have to be fulfilled independently and there is no coupling between the two dots. The potential of each dot is independent of the potential in the adjacent dot.

Special Case: Ruzin *et al.* If a capacitance C connects the dots to each other and to the leads, as indicated in figure 3-15, we should get the same result as Ruzin, Glazman *et al.* [89–91]. We are considering this as a special case to the formalism developed above. Appendix J derives the same formulas using Ruzin's approach. Since we want to compare the results to the formulas given by Ruzin, we adopt the same notation ($a \rightarrow 1$ and $b \rightarrow 2$). In this case we have

$$\begin{aligned} C_{ab} &= C \\ C_a &= C + C_1 \\ C_b &= C + C_2 \\ \tilde{C}^2 &= 3C^2 + 2C_1C + 2C_2C + C_1C_2 \ . \end{aligned} \tag{3.62}$$

For the charging energy we get

$$U = \frac{Q_1^2(2C+C_2)+Q_2^2(2C+C_1)+2Q_1Q_2C}{2\tilde{C}^2} + V_g(Q_1\frac{2CC_1+C_1C_2+CC_2}{\tilde{C}^2}+Q_2\frac{2CC_2+C_1C_2+CC_1}{\tilde{C}^2}) \quad . \tag{3.63}$$

Equation (3.63) should be identical to equations (8)-(10) in I. M. Ruzin's paper [89][1].

The two conditions for current flow result in the chemical potentials

$$\begin{aligned}\mu_1 &= \frac{e^2(N_1-\frac{1}{2})(2C+C_1)}{\tilde{C}^2}+\frac{e^2N_2C}{\tilde{C}^2}-V_ge\frac{(2CC_1+CC_2+C_1C_2)}{\tilde{C}^2} \\ \mu_2 &= \frac{e^2(N_2-\frac{1}{2})(2C+C_2)}{\tilde{C}^2}+\frac{e^2N_1C}{\tilde{C}^2}-V_ge\frac{(2CC_2+CC_1+C_1C_2)}{\tilde{C}^2} \quad ,\end{aligned} \tag{3.64}$$

and for the voltages, assuming the chemical potentials μ_1 and μ_2 are zero,

$$\begin{aligned}\frac{V_g}{e} &= (N_1-\frac{1}{2})\frac{2C+C_2}{2CC_1+CC_2+C_1C_2}+N_2\frac{C}{2CC_1+CC_2+C_1C_2} \\ \frac{V_g}{e} &= (N_2-\frac{1}{2})\frac{2C+C_1}{2CC_2+CC_1+C_1C_2}+N_1\frac{C}{2CC_2+CC_1+C_1C_2} \quad .\end{aligned} \tag{3.65}$$

A comparison with Appendix J shows that the two derivations have identical results.

Special Case: Capacitance Between Dots In this case we neglect all capacitances except the capacitances from the gate to the dots and the coupling capacitance, as shown in Fig. 3-16.

[1]There are two errors in I. M. Ruzin's paper, the first one is a sign error in eq. (6), the second an indexing error in eq. (9).

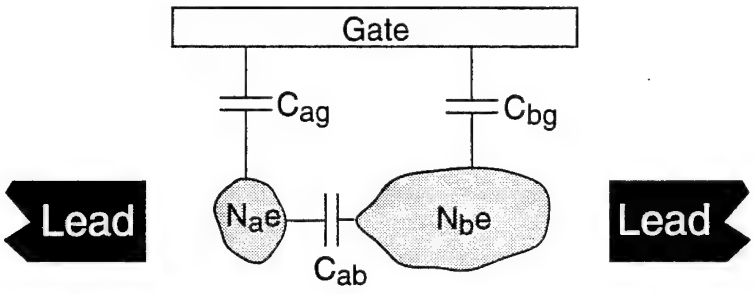


Figure 3-16: Easiest case for coupling of two quantum dots.

For the capacitances in this system we get

$$\begin{aligned} C_a &= C_{ag} \\ C_b &= C_{bg} \\ \tilde{C}^2 &= C_{ab}C_{ag} + C_{ab}C_{bg} + C_{ag}C_{bg} \ . \end{aligned} \tag{3.66}$$

For the charging energy we get

$$U = \frac{Q_a^2(C_{ab}+C_{bg}) + Q_b^2(C_{ab}+C_{ag}) + 2Q_aQ_bC_{ab}}{2(C_{ab}C_{ag}+C_{ab}C_{bg}+C_{ag}C_{bg})} + V_g(Q_a+Q_b) \ . \tag{3.67}$$

This is identical to eq. (J.6), which was derived using Ruzin's formalism.

For the chemical potential in the dots and the conditions for conduction we get

$$\begin{aligned} \mu_a(N_a,N_b) &= \frac{e^2(N_a-\frac{1}{2})(C_{ab}+C_{bg}) + e^2N_bC_{ab}}{C_{ab}C_{ag}+C_{ab}C_{bg}+C_{ag}C_{bg}} - eV_g \\ \mu_b(N_a,N_b) &= \frac{e^2(N_b-\frac{1}{2})(C_{ab}+C_{ag}) + e^2N_aC_{ab}}{C_{ab}C_{ag}+C_{ab}C_{bg}+C_{ag}C_{bg}} - eV_g \end{aligned} \tag{3.68}$$

$$\begin{aligned} \frac{V_g}{e} &= \frac{(C_{ab}+C_{bg})(N_a-\frac{1}{2}) + C_{ab}N_b}{C_{ab}C_{ag}+C_{ab}C_{bg}+C_{ag}C_{bg}} \\ \frac{V_g}{e} &= \frac{(C_{ab}+C_{ag})(N_b-\frac{1}{2}) + C_{ab}N_a}{C_{ab}C_{ag}+C_{ab}C_{bg}+C_{ag}C_{bg}} \ . \end{aligned} \tag{3.69}$$

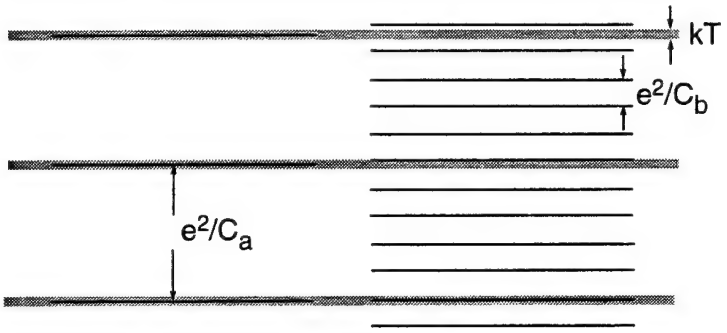


Figure 3-17: Energy ladders for a double dot system. The left ladder shows the smaller dot, the right the larger dot. Note that the two ladders move with different speed when the gate voltage is changed, according to the difference in capacitive lever arms.

3.5.1 Position of Conductance Peaks

A double dot system can only conduct current if the chemical potentials of both dots are aligned with the chemical potentials in the leads. A simple picture of the problem is shown in Fig. 3-17. In the case of no coupling between the quantum dots we have two rigid energy scales which move with respect to each other, but whose movement is only dependent on the gate electrode rather than the charge on the adjacent dot. Conduction through the double dot can be sustained if the chemical potentials, plotted here, are within about k_BT of each other and of the leads.

Because the levels move with respect to each other, it is not easy to extract information on the location of conductance maxima from Fig. 3-17. Instead, one can plot a diagram of charging energy as a function of gate voltage, as done in Fig. 3-4 for the single dot. We use eq. (3.53) to plot the charging energy as a function of gate voltage for the double dot system. The charging energy has to be plotted with both N_a and N_b as parameters. For identical dots and no coupling capacitance, the states of lowest energy have identical numbers of electrons in each dots, as shown in Fig. 3-18. As can be seen in Fig. 3-19, an unequal number of electrons in the dots is only possible at the charge degeneracy points. The energy levels for (2,3) and (3,2) are degenerate for any gate voltage. At the intersection of (2,2) and (3,3), current

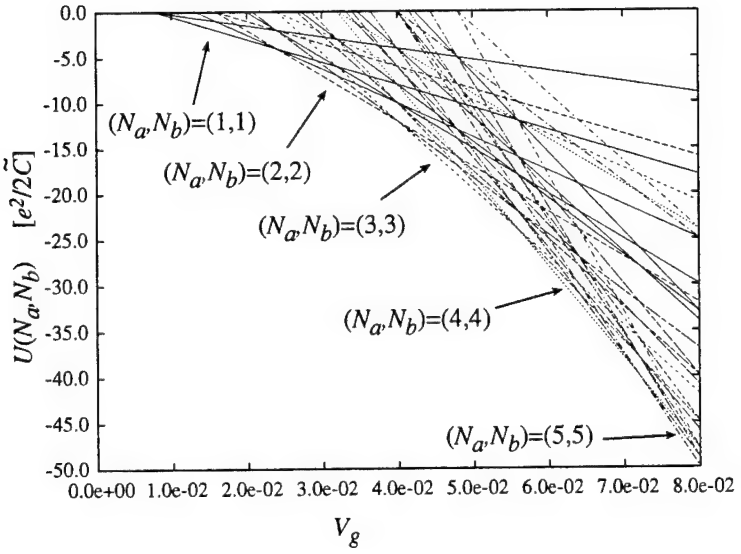


Figure 3-18: Charging energy as a function of gate voltage for two identical dots, with $C_{ag} = C_{bg} = 10\,\mathrm{aF}$.

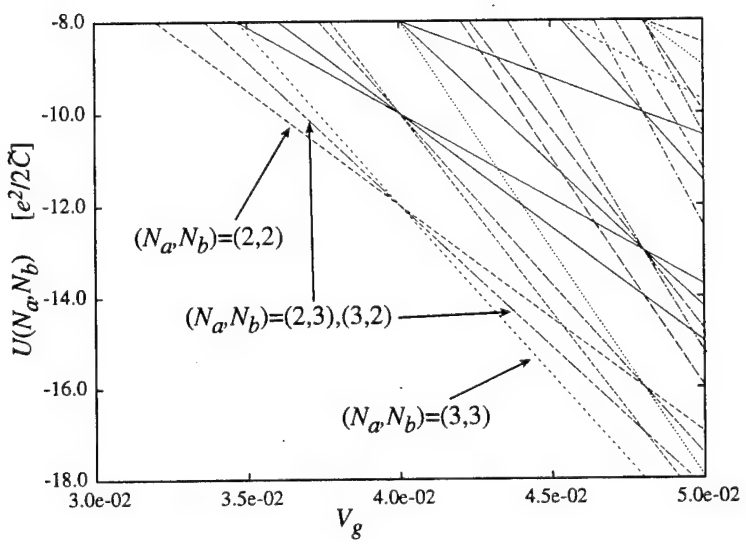


Figure 3-19: Close-up of Fig. 3-18. The energy levels that contribute to current flow are labeled.

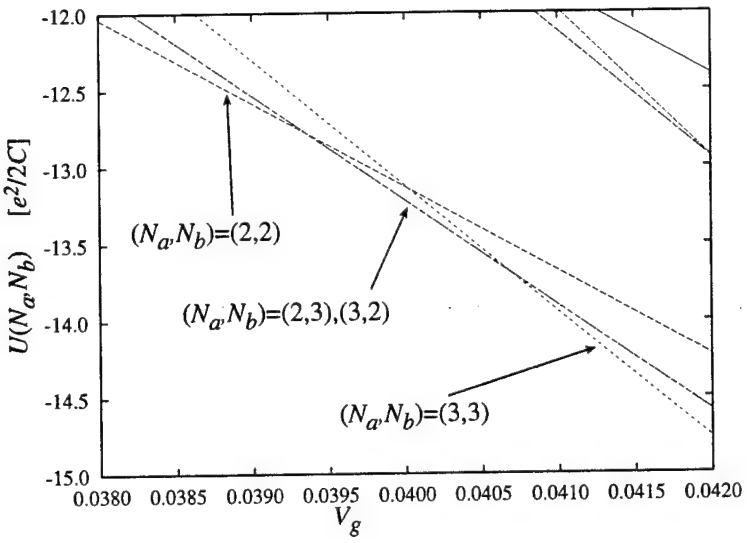


Figure 3-20: Close-up of charging energy as a function of gate voltage for two identical dots, in case of coupling between the dots.

can flow through the system, while the system goes through the two possible cycles: (2,2) → (3,2) → (2,3) → (2,2), or (3,3) → (3,2) → (2,3) → (3,3). We have assumed that the stable states are (2,2) and (3,3) and that electrons tunnel from the left to the right. If a coupling capacitance, in this case of 1 aF, is introduced into the problem, the energy levels (2,3) and (3,2) are still degenerate, but shifted to a slightly lower energy, as shown in Fig. 3-20. This causes an intermediate state, (2,3) and (3,2), to be most favorable for a small range of gate voltages. For the case of coupling, the conductance peak splits into two, the first one with the cycle involving the state (2,2), and the second peak involving the cycle with the state (3,3). This splitting was recently observed experimentally by Waugh *et al.* [94]. If there is a detuning between the dot-to-gate capacitances, the degeneracy of the (2,3) and (3,2) level is lifted, as shown in fig 3-21. For conductance to occur in this case, the electron would have to be thermally-excited because the state (3,2) is necessary for conduction through the double dot system.

If the gate capacitances strongly deviate from each other, which is the case for unequal dots, the number of electrons added for a sweep of the gate voltage is different

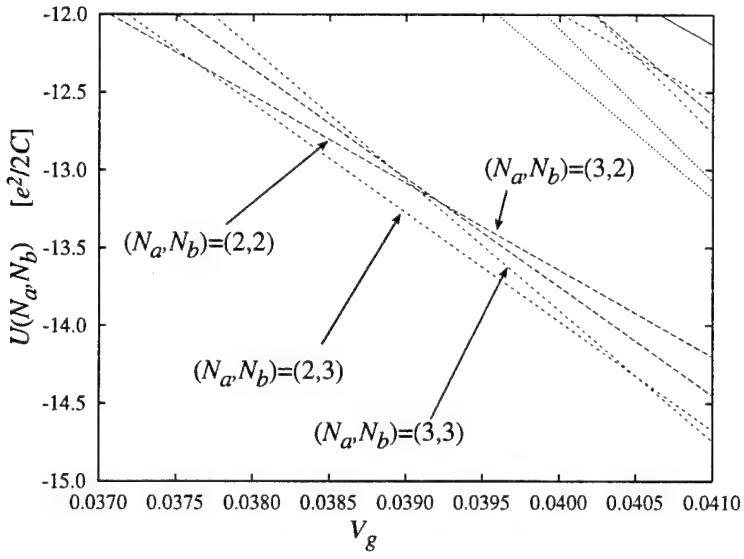


Figure 3-21: Close-up of charging energy as a function of gate voltage for two identical dots, in case of coupling and 5% detuning between the dots.

for each dot. For the case $C_{ag} = 10\,\mathrm{aF}$, $C_{bg} = 5C_{ag}$ and $C_{i\Sigma} = 10C_{ig}$, the charging energy as a function of gate voltage is plotted in Fig. 3-22. As expected, for one electron added to the smaller dot, five electrons have to be added to the larger dot. A close-up of the transition from the state $(N_a, N_b) = (1, 7)$ to $(2, 8)$ is shown in Fig. 3-23. Because we chose the ratio of the capacitances to be an integer, the states with $(1, 7)$ and $(2, 8)$ happen to be degenerate. The conduction mechanism is identical to the one discussed for identical dots. For slightly different choices of capacitances, one of the two states would be lower in energy and the other state could be reached only through thermal excitation. Both states, (1,8) and (2,7), are necessary for conduction.

Using the values of capacitance measured for the single dots given in Table 3.1, we can plot the charging energy as a function of gate voltage, shown in Fig. 3-24. The addition of an electron to the smaller dot does not coincide with the addition to the larger dot. In this case, we expect the conductance to be negligible and only thermally-activated. A close-up of the transition from (1,5)→(1,6) →(2,6)→(2,7) is shown in Fig. 3-25. At the charge degeneracy (1,6),(2,6) current can only flow if the intermediate states (1,5) and (2,5) are also accessible. At this gate voltage,

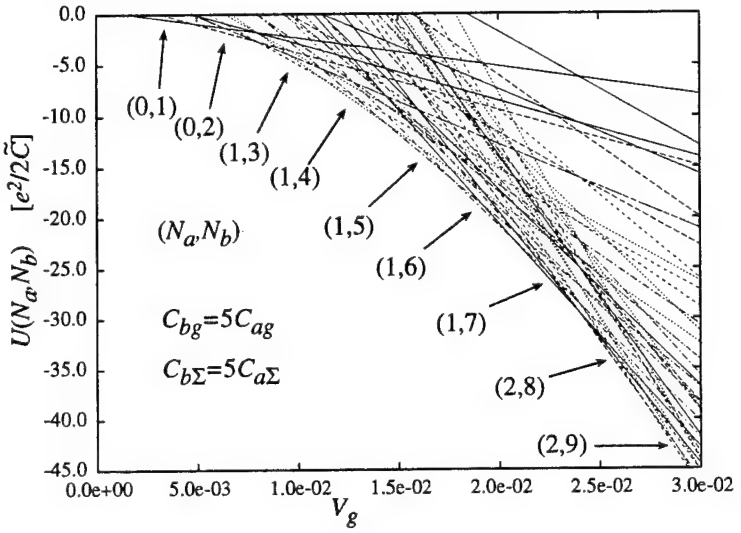


Figure 3-22: Charging energy as a function of gate voltage for two unequal dots, with $C_{bg} = 5C_{ag}$, $C_{b\Sigma} = 5C_{a\Sigma}$, and $C_{ag} = 10\,\mathrm{aF}$.

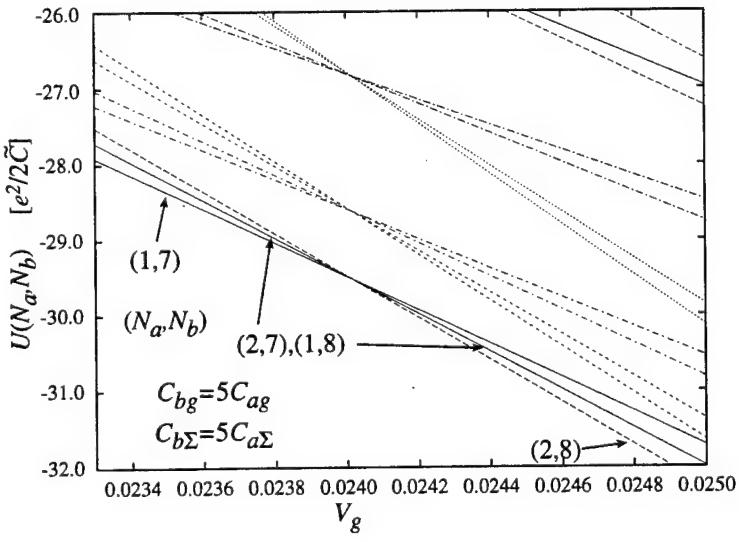


Figure 3-23: Close-up of charging energy as a function of gate voltage for two unequal dots.

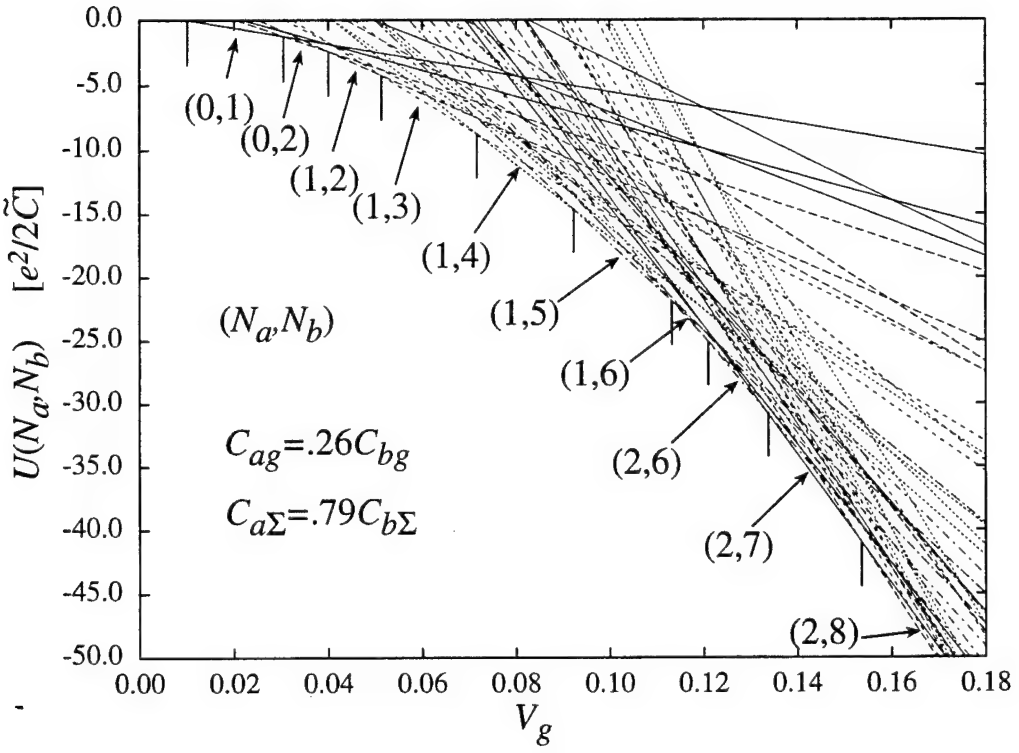


Figure 3-24: Charging energy as a function of gate voltage for two unequal dots using capacitances measured for separate dots given in Sec. 3.4.

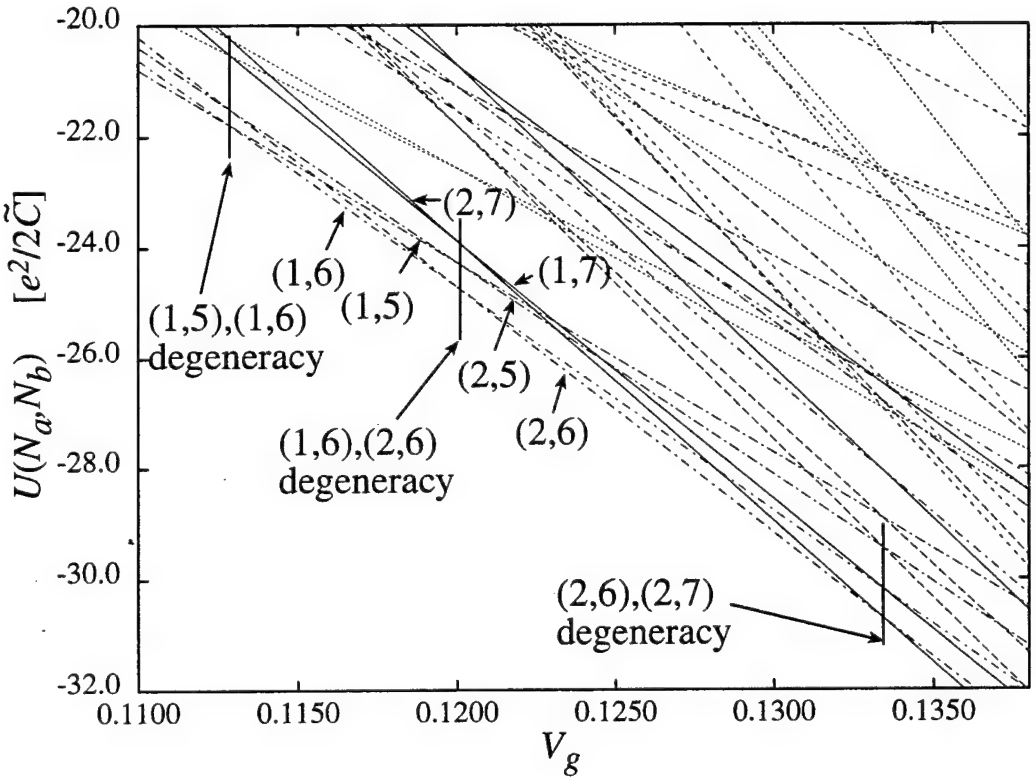


Figure 3-25: Close-up of charging energy as a function of gate voltage for two unequal dots using capacitances measured for separate dots given in Sec. 3.4.

these states happen to be degenerate. They are not lowest in energy, but still lower than (1,7) and (2,7), which could also provide a conduction mechanism. If the initial state were (1,6), current could flow through the cycle (1,6)→(2,6)→(2,5)→(1,6) or (1,6)→(1,5)→(2,5)→(1,6). It should be noted, however, that conduction is more likely at the degeneracy point of (1,5) and (1,6) since the energy levels (2,5) and (2,6) are closer at this point. The thermal energy, using a temperature of 300 mK, is about $0.1\,e^2/2\tilde{C}$. Considering the scale of Fig. 3-25, conduction through the double dot system is unlikely at this temperature. The energy difference from the charge degeneracy point (1,6) and (2,6) to the levels (1,5) and (2,5) is approximately $5\,k_BT$. One would have to look into events where the addition of an electron to the smaller dot is closer in gate voltage to the addition of an electron to the larger dot. Ideally, this would be a point of multiple degeneracy, such as the one shown in Fig. 3-23. If the ratio between the gate capacitances is a rational number, such an event can occur regularly as a function of gate voltage. If the ratio is not a rational number, we expect to observe a stochastic Coulomb blockade effect, as predicted by Ruzin [89] and observed by Kemerink [95].

Another representation of the conductance condition emerges if one plots the conditions on the chemical potential for the independent dots, described in eq. (3.60). We neglect all electrodes except one gate, and get for the chemical potential

$$\mu_a = \frac{e^2}{C_{a\Sigma}}(N_a - \frac{1}{2}) - eV_g\frac{C_{ag}}{C_{a\Sigma}} \quad . \tag{3.70}$$

If we consider only one dot, as shown in Fig. 3-26, we can plot the (negative) chemical potential versus gate voltage. If electrons can tunnel into and out of the dot, the chemical potential as a function of gate potential is a sawtooth function. For every discontinuity, N_a is increased by one. If we plot instead $-\mu_a + e^2(\hat{N}_a - 1/2)/C_{a\Sigma}$ versus gate voltage, where $\hat{N}_a$ increments by one at each charge degeneracy, we get a straight line with slope $eC_{ag}/C_{a\Sigma}$ as a function of gate voltage. The factor $C_{ag}/C_{a\Sigma}$ is again the capacitive lever-arm. This straight line can be collapsed onto the y-axis without

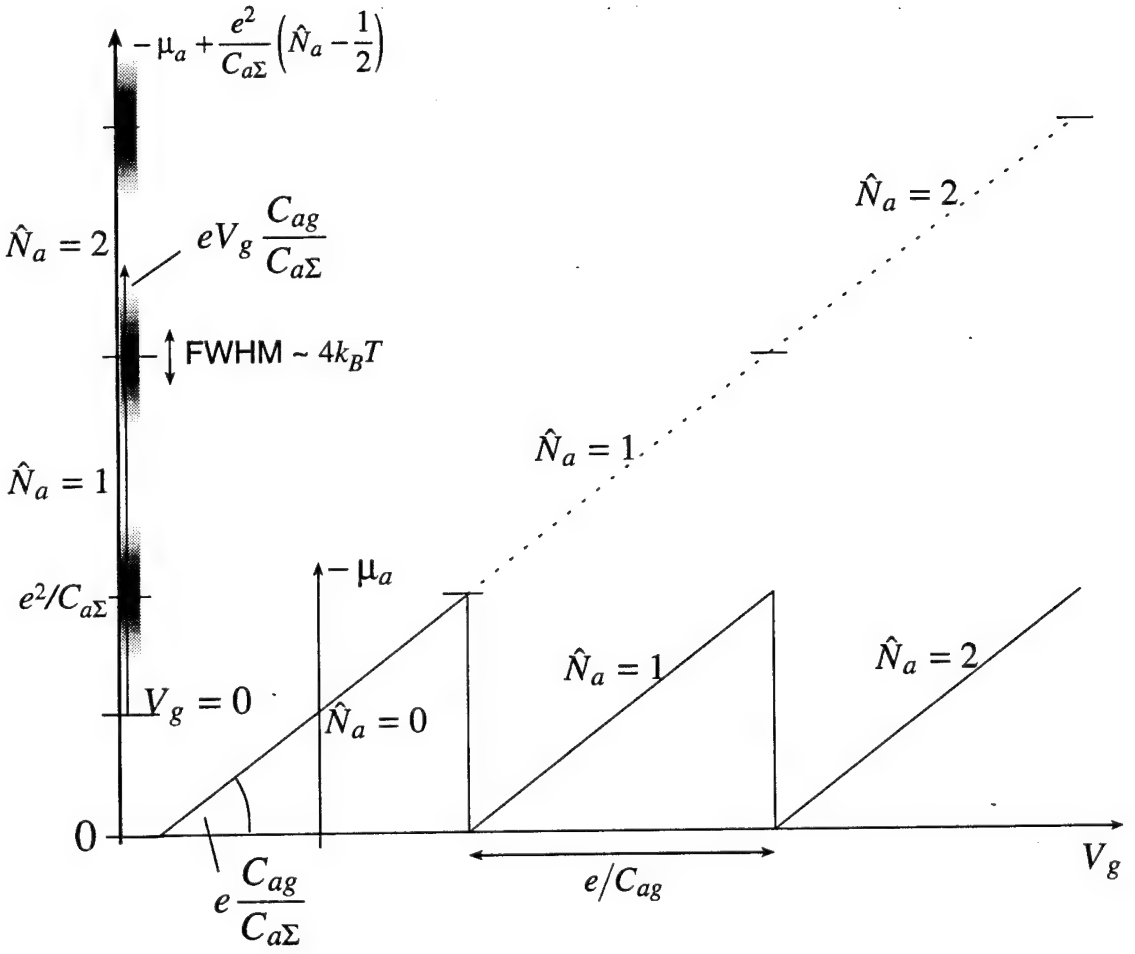


Figure 3-26: Graphical solution for eq. (3.60), for the case of only one dot.

loss of information. The gate voltage is parallel to the axis, moving with a speed of $-eV_gC_{ag}/C_{a\Sigma}$, and starting offset from the origin by $e^2/2C_{a\Sigma}$. The conductance peaks at every charge degeneracy point, when $\mu_a = 0$, and is measurable a few k_BT around these points.

For two dots, this graph can be extended to include a second dot on the other axis, as shown in Fig. 3-27. The x- and y-axes represent a measure of the chemical potential for dots a and b, $-\mu_a+e^2/C_{a\Sigma}(\hat{N}_a-1/2)$ and $-\mu_b+e^2/C_{b\Sigma}(\hat{N}_b-1/2)$, where $\hat{N}_a$ and $\hat{N}_b$ are the number of electrons in the dots in equilibrium. These numbers are changing every time that an increase in gate voltage causes a charge degeneracy to happen, represented through vertical and horizontal lines. The gate voltage changes the chemical potential in the dots, indicated by the line in the plot. The capacitances between the gate and the dots cause the chemical potentials in the two dots to sweep with different speed. The slope of the line is given by ratio of the capacitive lever-arms of the gate to the dots. Conductance peaks appear when the line of the gate voltage intersects the lines of charge degeneracy for both dots simultaneously. The circles around these resonance points represent the thermal spread of a few k_BT around each charge degeneracy point, as given by eq. (3.38). Figure 3-27 is a picture of a system with $C_{bg} \simeq 4C_{ag}$ and $C_{b\Sigma} \gtrsim C_{a\Sigma}$. This represents approximately the conditions for the charging energy versus gate voltage, plotted in Fig. 3-24. The ratio between the gate capacitances causes about four electrons to be added to the larger dot before one electron can be added to the smaller dot. A few of the electron numbers in the dots are indicated in the figure. If there were no coupling from dot b to the gate we would obtain the situation depicted in Fig. 3-26 if conductance through the inactive dot b is possible, or no conduction at all, if dot b is blockaded. A close-up of the chemical potentials at the gate voltage where the second electron is added to the smaller dot is shown in Fig. 3-28. In this close-up, the gate voltage range is chosen similar to Fig. 3-25, and the three charge-degeneracy points are indicated by the black dots. The double-dot can only conduct if the charge degeneracy point is close to a common

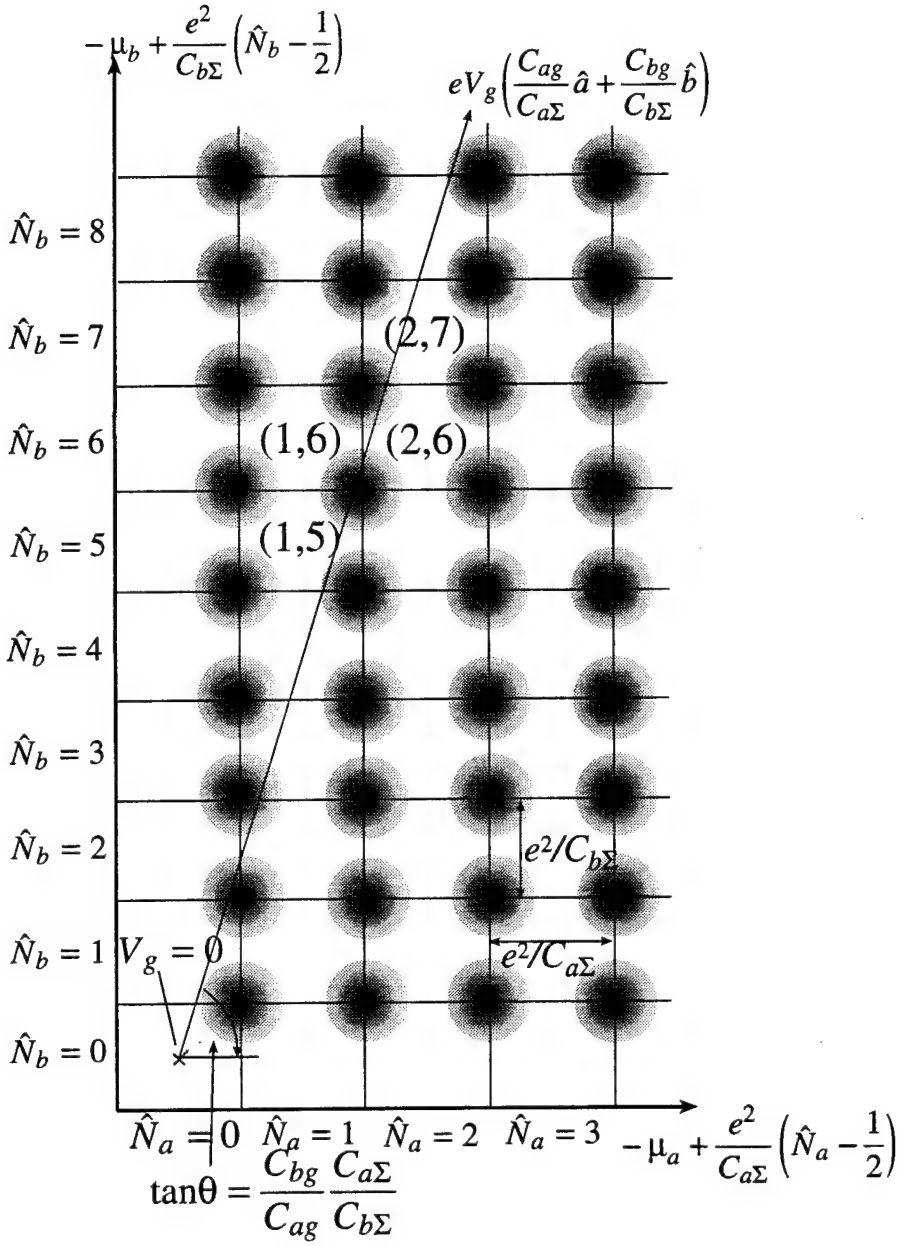


Figure 3-27: Plot representing change in chemical potentials in both dots. Horizontal lines represent charge degeneracy in dot b, vertical lines in dot a. The slope of the line representing the change of chemical potential due to gate voltage change depends on the ratio of the capacitive leverarms of the dots. Current can flow if four adjacent charge states are accessible within a few k_BT.

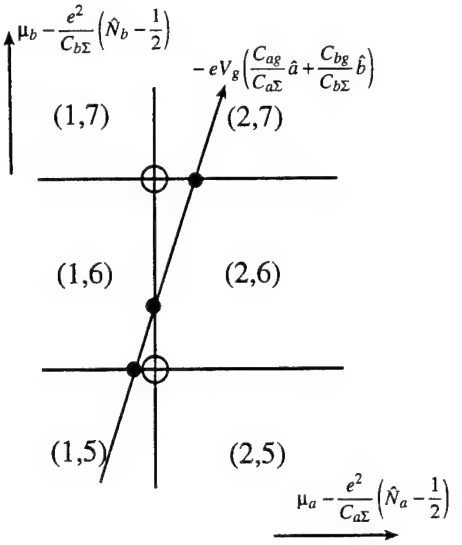


Figure 3-28: Close-up of Fig. 3-27 for a similar gate voltage range as in Fig. 3-25. Current can only flow if all states adjacent to the circles are accessible. The filled dots represent the charge-degeneracy points.

charge degeneracy point of both dots, indicated as circles. At the (1,5)-(1,6) charge degeneracy point, the system can conduct only if the location of the degeneracy is within a few k_BT of the circle, indicating charge degeneracy of (1,5), (1,6), (2,5), and (2,6).

3.6 Measurement of a Double Dot

The device shown in Fig. 3-7 was biased in a configuration shown in Fig. 3-14. The conductances, measured as a function of gate voltage along with conductances of the individual dots for comparison, are shown in Fig. 3-29. A modulation of two frequencies can be observed for the double dot case. The two frequencies appear similar to the frequencies observed for the 200 and 400 nm quantum dots. An analysis of the spacing between conductance peaks reveals that the capacitances are larger than for the original dots. The capacitances for the individual dots and the coupled dots, extracted from this plot, are shown in Table 3.2. The first tall peak for the double dot system can be fit using eq. (3.38), similar to the fits for the individual

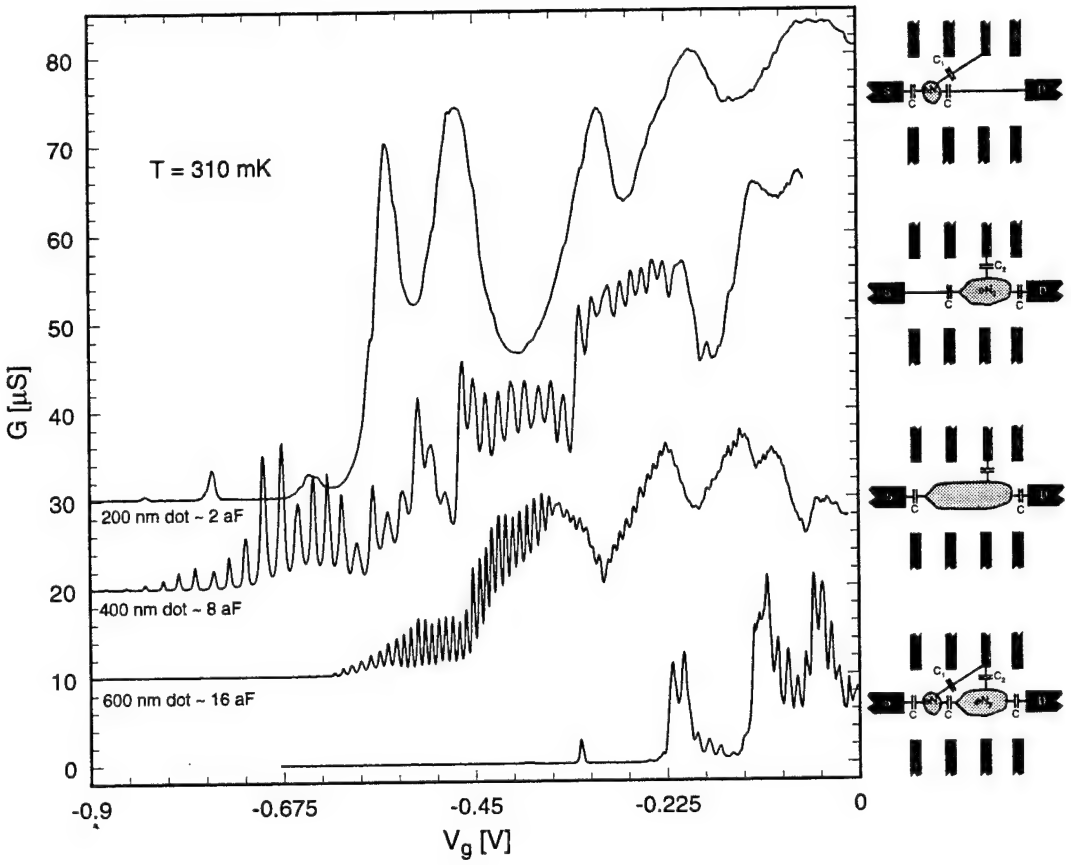


Figure 3-29: Conductance as a function of gate voltage for separate dots of 200, and 400 nm length and a 200 and 400 nm dot in series. For comparison, the conductance for the 600 nm long dot is plotted as well.

	200 nm (*a*)	400 nm (*b*)
individual dots [aF]	2.0	8.9
coupled dots [aF]	2.6	11.8

Table 3.2: Capacitances for 200 nm dot (a) and 400 nm dot (b) for the dots measured individually and as coupled dots.

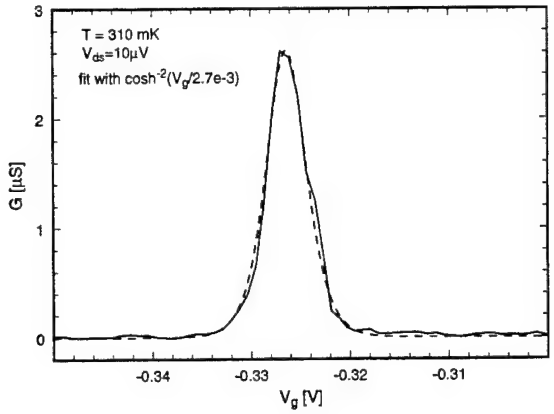


Figure 3-30: Fit of the first conductance peak in the double dot case with the Fermi function.

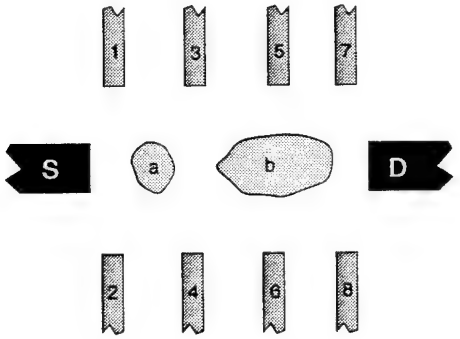


Figure 3-31: Nomenclature for the electrodes in the experiment.

dots in Fig. 3-13. The result is shown in Fig. 3-30. The conductance peak fits the same temperature and lever-arm as the 400 nm dot alone.

3.6.1 Threshold Voltage Shift

The threshold voltage for the onset of conduction is significantly higher for the double dot system than for individual dots. To discuss this, we adopt the nomenclature shown in Fig. 3-31 for the electrodes. For the explanation of the threshold shift, we consider the conductance peaks of the smaller dot, which are easier to identify due to their sparsity. The gate electrode 5 was swept in all experiments. The capacitance C_{a5} sets the distance between conductance peaks. If the potential on any of the electrodes i

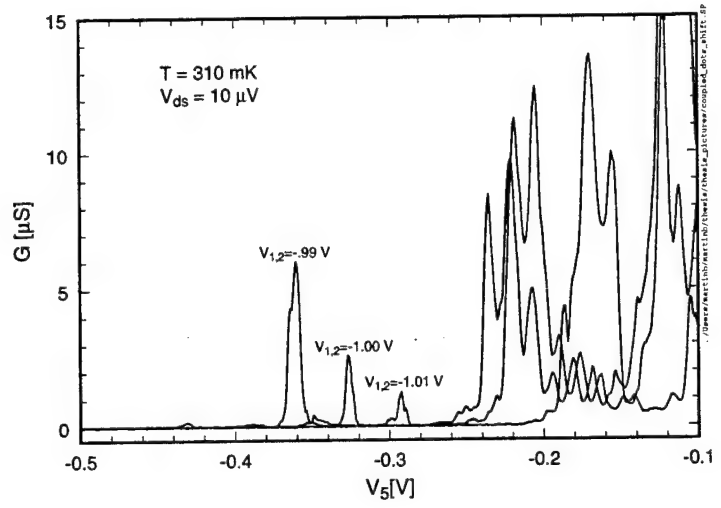


Figure 3-32: Shift of the peak associated with the smaller dot.

is changed by ΔV_i, each conductance peak is shifted by

$$\Delta V_g = \frac{C_{ai}}{C_{ag}} \Delta V_i \ , \tag{3.71}$$

according to eq. (3.61). The shift of a known peak reveals the capacitance of the electrode to the dot. Figure 3-32 shows the first strong conductance peak for three different potentials of electrodes 1 and 2, changed by 10 mV increments. Using eq. 3.71, the capacitances C_{a1} and C_{a2} are approximately 3.4 aF. Using the same method, we can extract capacitances for electrodes 7 and 8 to be $C_{a7} \simeq C_{a8} \simeq 1.1$ aF. These are the electrodes that have to be changed if one wants to bias the large dot in addition to the already existing small dot. To create a third tunneling barrier, gates 7 and 8 have to be swept negative, changing the gate voltage at which a fixed number of electrons will be in the dot. This is the reason for the threshold voltage shift in Fig. 3-29.

Chapter 4

Conclusion

4.1 Summary

This report has explored two critical aspects of future electronics. It has described fabrication technology for nanometer-scale semiconductor devices and measurements of novel quantum dot devices that were made using this technology.

In order to make devices which use lateral tunneling structures, lithography has to be pushed into the sub-100 nm region. Lithographic definition of tunneling barriers has to be on the order of a few tens of nanometers and control over these dimensions has to be even more stringent. At these dimensions, the lithographic task to build lateral tunneling devices is more demanding with regards to resolution than other applications of x-ray lithography, such as MOS gate levels or optoelectronic devices. These requirements leave x-ray and e-beam lithography as the only possible tools for the definition of the gate level. E-beam lithography is a process featuring high resolution and high flexibility in pattern formation. However, it is not a technique which can be used in large-scale fabrication processes because of its serial nature. Other advantages of x-ray lithography, directed especially towards the engineering community, include superior process latitude, printing of high aspect ratio resist lines, and the ability to write large area patterns on x-ray masks. E-beam lithography can

not easily be used to write large area coherent structures because of the stitching error at field boundaries. This work used e-beam lithography only to lay down a pattern on a parent x-ray mask, which was then replicated once onto a daughter x-ray mask and then printed onto a substrate using x-ray lithography. In the course of this work, new technologies were developed for the replication and processing of x-ray masks. Apart from the diffusive wires, made by John Scott-Thomas, which initiated the discovery of Coulomb blockade in semiconductor structures, these experiments marked the first time that Coulomb blockade devices were fabricated using x-ray lithography.

A unique device consisting of four adjacent quantum point contacts was built and measured at cryogenic temperatures. In the course of this work, the measurement electronics and data acquisition system were designed and built. The device was measured at a temperature of 300 mK in a commercially-available probe. The design of the device allowed biasing of quantum dots of various sizes, as well as biasing of a double dot structure. Using two adjacent quantum point contacts to bias the device, the center-to-center distance of tunneling barriers was 200 nm for the smallest dot and 600 nm for the largest. Conductance oscillations as a function of gate voltage were seen for quantum dots of all sizes and the capacitances of these dots were extracted. The gate-to-dot capacitance of the smallest quantum dot was of the order of only 2 aF. The device was also biased to produce two consecutive quantum dots of 400 and 200 nm length, as measured from the center of the quantum point contacts defining the dot. A beating pattern of two different frequencies was observed for the conductance oscillations as a function of gate voltage. For zero temperature, quasi-random conduction resonances would be expected. Because of the relatively high temperature during the measurements, the conductances peaks are broadened, leading to the observed beating pattern.

4.2 Future Work

This report has shown that it is possible to make quantum-effect devices using technology that is suitable for mass fabrication. It has investigated the transition from single to coupled quantum dots. A logical continuation is the investigation of more than two quantum dots in close proximity, leading eventually to large arrays. This effort could merge previous experiments, which employed grid gates to produce periodic potentials within a 2DEG[7,8,14,96], with more recent work on coupled quantum dots[40,78,94]. So far, most of the work on quantum dot devices have employed only one high-resoution lithography step, the gate level. With increasing device density, it will be necessary to integrate other lithography steps, most likely the mesa etch, and possibly the ohmic contact level, into the fabrication process. For this to work, high resolution alignment methods have to be used. All of these technologies have been developed in the NanoStructures Laboratory, but so far have not been integrated into one whole process. Although the gate level will most likely require PMMA as resist to ensure high resolution, mesa and ohmic contacts could employ chemically-amplified resists instead, reducing exposure times and making the process more compatible with manufacturing environments. Another promising area is to try to diminish dot capacitances by using new material systems such as SOI which are specifically designed to reduce parasitic capacitances [83]. Coulomb blockade could potentially be observed at liquid nitrogen temperatures, making the issue of single-electronics much more plausible to the engineering community. This effort could be based in part on the expertise of the SOI research groups at MIT.

Appendix A

Mask Fabrication Process

The process of patterning, developing, and electroplating an x-ray mask is at the moment done by graduate students. Ultimately, as the process becomes more common and ceases to be an area of research interest, this task should be transferred to a technician. This process outline describes the steps for the fabrication of an x-ray mask by a technician under supervision of a graduate student.

An approximate time schedule is shown in Fig. E-1. Depicted is a schedule for the fabrication of an x-ray mask which could be finished within about one working day. Note that the times noted are estimates in case no major problems appear during the fabrication sequence. It is possible to cut down on time, mostly during the pump-down for the RIE and SEM. It should be noted, however, that because of the parallelism of the processes this might not decrease the total time by too much if only one person is involved in the process. Characterization of the plating bath and mixing of the developer for example has to be done during the first RIE pump-down, no matter how long that pump-down takes.

- Turn on the pump for the plating system and the temperature controller for the water bath. The plating system needs a while to stabilize to the desired bath temperature (33 °C). The set point for the water bath might have to be adjusted to reach the desired bath temperature. It is generally around 29 °C.

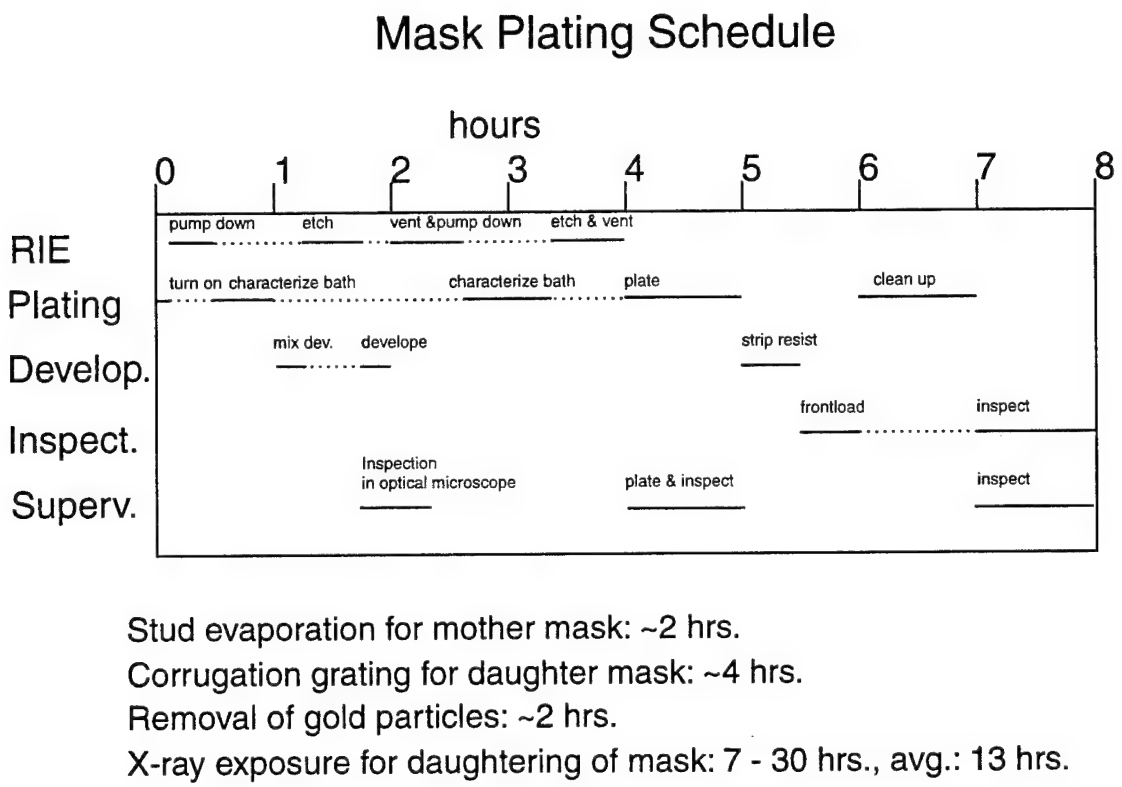


Figure A-1: Approximate lab schedule for the mask fabrication process. To make a working daughter, it is necessary to run through this process twice. Extra time must be devoted for the evaporation of studs for the mother mask, fabrication of the corrugation grating in case the daughter mask is to be used in intimate contact with the substrate, the removal of gold particles by wet etching, and the x-ray exposure for daughtering of the mask.

- Start a clean-up run on the RIE to set up the parameters for an oxygen de-scum etch on the Plasma-therm RIE. Etch parameters right now are He (10 sccm) O_2 (2.5 sccm) at 6 mTorr for 12 sec. Power settings are 44 W at 75 V DC. He backside cooling is set to 20 °C at 7 Torr and 2.6 sccm He flow. This step should only be necessary if the same RIE was used previously with different gases.

- During the clean-up run, the plating bath can be characterized. This involves measurements of pH-value and conductivity. In some cases an IV-characterization might be advisable. Brightener should be added if needed.

- Mix the developer (1:2 MIBK:IPA) in a 1 liter beaker. The desired temperature for the developer of 21 °C can be achieved by heating of the developer on a hotplate or warming it with the hands on the outside of the beaker. The latter method avoids an overshoot of the temperature. In case the temperature is too high, the developer can be cooled down by spraying IPA on the outside of the beaker and blowing with the nitrogen gun. Monitor the temperature of the developer. The temperature should drop to about 16 °C upon mixing. If that doesn't happen something is wrong with the chemicals used.

- Once the developer temperature has settled to 21 °C, the mask is developed for 90 seconds by slow and constant agitation in the developer. The development should be kept the same for all masks.

- When the development time is up, immediately rinse the mask with IPA using the squirt bottle.

- Blow the sample dry with the filtered N_2 gun. It is critical not to blow on the sample too hard. The pressure on the N_2 gun can be reduced. Make sure to blow dry starting in the middle of the mask and work your way outside.

- optically inspect the mask with an optical microscope in clearfield and darkfield. This should be done by a person familiar with the pattern.

- Load the sample into the RIE for de-scum.

- Etch for approx. 5 seconds using the parameters from the clean up run saved in the computer.

- Keep the sample under vacuum in the RIE until you are ready to plate.

- Electroplate 200 nm of gold. This should be done by a person familiar with the pattern. Plating thickness should be determined using optical transmission, but other methods such as alpha-step can be used as well.

- Check the pattern by inspection in the optical microscope and, if needed, the scanning electron microscope. The inspection should be performed by a person familiar with the pattern.

- Strip the resist by rinsing in acetone and methanol and plasma ashing. Alternatively, the resist can be stripped by soaking the mask in hot NMP (~90°C).

- Check on an optical flat if there are any large particles on the mesa that prevents a uniform gap. Alternatively, the mask can be brought in close proximity with another mask, as if a daughter exposure was set up. The process for removal of the particles is to spin photo resist over the whole mask and to expose the suspicious particles for about 2 minutes in the optical microscope with maximum white light intensity. Then use an eye dropper to put drops of gold etch on the openings in the resist. The etch of the largest particles should be completed in about 20 minutes.

- For masks to be used in microgap exposures, evaporate Al-studs. The height of the studs should be about 3 μm. For mask replication, the Al-studs are not always necessary because the natural warp of the mesa sets an initial gap of about 3 μm.

- Do a daughter exposure. It is typically advisable to do a first exposure in head 1 with a gap determined by the studs. If the effects of diffraction prevent successful replication, another exposure can be made in head 3 using the vacuum fixture and a gap of about 1 μm. The gap can be held constant using electronic feedback (see Appendix I).

- Flood expose the daughter mask using deep UV. A TEM grid and a thin wire should be used to monitor development and plating. Typical exposure time is 30 min.

- Process the daughter mask. The processing sequence is the same as for the mother mask, only that the development time is not fixed and should be done by a person familiar with the pattern. Also, puddle development rather than immersion development is used for the daughter mask because of improved particle control. The improved temperature control of immersion development is not necessary for daughter masks because of the larger process latitude offered by x-ray exposures compared to e-beam exposures.

- If the daughter mask is to be used for exposures in intimate contact with the substrate, a polyimide corrugation grating is necessary for outgassing of the resist. For these steps it is a good idea to process a silicon monitor in parallel:

  - Strip PMMA from exposed daughter mask. One good method of stripping is to let the mask soak in hot NMP (90°C) for 15 min. Be careful: NMP has a flash point of 93 °C. A temperature controlled hot plate should be used.

  - If the plating base is to be removed in the alignment mark regions: spin photoresist, e. g. Shipley 1813.

  - Expose the alignment mark regions for appropriate time. I typically use the substage condenser and expose through the membrane which causes the

thick gold areas to remain covered with resist (no loss of contrast during gold etch!). Typical exposure times I use is 10 sec. with the substage condenser aperture fully open and the Leitz Xenon light as the light source.

- Develop with Shipley 351 (1:5 351:DI) developer for about 30 sec.
- Plasma ash the sample in He/O_2 for 6 seconds at about 50 W power. This step is absolutely critical to get etching in all areas at the same time with good uniformity. I didn't get any consistent results without this step. This step is probably necessary because of the standing wave effect in optical lithography. An intensity node can appear slightly above the surface of the sample and cause a very fine resist layer to remain. It might be possible to cut down on the exposure time though. The time should not be increased, however, because prolonged exposure to the plasma crosslinks the photoresist which makes it difficult to remove.
- Etch the gold in 10:1 dilute gold etchant for 20 seconds.
- Etch the NiCr in 10:1 dilute NiCr etchant for 5 seconds. This etch is much more dangerous than the gold etch because it tends to etch laterally at an incredible speed and undercuts everything, causing Moire alignment marks, for example, to fall over easily.
- Spin about 300 nm of PI, use 3:2 PI2610:NMP at 3 krpm for 60 sec.
- Clean off the polyimide at the edges of the mask with a soaked Q-tip for good electrical contact with the mask holder. Remember: we use electrostatic contact, so we have to make electrical contact to the mask.
- If the polyimide should be stripped in the alignment mark area, bake this layer at 140 °C like the following layer. Spin photoresist, develop, strip, then hardbake the polyimide. If polyimide should not be etched away, just bake at 130 °C and 250 °C for 30 minutes each.
- Spin on 5:15 PI:NMP to get roughly 70 nm of polyimide.

- Bake at 140 °C for 30 min. This temperature is critical.
- Immediately spin on photoresist and bake for 30 minutes.
- Expose with a 100 μm-period grating flex mask, and if necessary, expose the alignment marks.
- Develop for 30 seconds using 5:1 DI water:Shipley 351 developer. Use the monitor before you do your mask.
- Examine the grating in the optical microscope.
- Strip the remaining resist using acetone and methanol.
- Bake at 200 °C for 1 hour.

Appendix B

Layout Transfer from MIT to NRL

After the device layout is finished, and after proper alignment marks and diagnostics patterns, as described in Appendices F and E, are dropped in, the KIC file needs to be converted into JEOL readable format and transferred to the e-beam site, in this case NRL.

- If the KIC layout file was designed on a machine other than the NeXT machines transfer the file via ftp to the NeXT account. Every directory that is used for the layout contains a file for the definition of the layers, called '.KIC'. This file should be sent along with the layout file. You'll need it for the kictocif conversion.

- Convert the file from KIC format to CIF format using the command kictocif. This command should be run on the NeXT machines. For some reason other computers such as MTL have problems handling the large integers used for x-ray masks (remember: our unit length λ is usually only 10 nm, which is .01 microns per lambda. This causes large integers to appear in the design of a 1×1 cm die.).

- Ftp the CIF file to Bill Chu's account on submicron.ll.mit.edu. Ask somebody for the password if you don't know it. Put your file in a subdirectory with your name to avoid confusion with all the .KIC and setup.cnv files from other users.

- Quit ftp and start X-windows on the machine you want to use for the conversion.

- Open the display for submicron.ll by typing "xhost + submicron.ll".

- Remote login into submicron.ll.mit.edu, selecting X-windows at the first prompt and then typing your machine name. Then go to your directory.

- Run the conversion program by typing cnv from the shell. If you have a setup.cnv program in your directory already, it will ask you if you want to use the existing one. If your .KIC file didn't change from last time and you didn't change part of your layer to single pass lines, you can use the old setup.cnv. Otherwise you have to edit the existing one.

- Use the arrow keys to go from one item to the next. Use the space bar to change a selection. If you edit your setup.cnv file, here is what you have to set for each separate layer: you want to use the 5th lens rather than the 4th lens. The field size should be 80 μm instead of 100 μm. If you use single pass lines in one particular layer, line rank should be different from shot rank. The origin of the field should be set manually to 0 μm/0 μm. This is the number you want to change if you have problems concerning field boundaries going through your fine features. After you are done, select "Continue".

- In the next menu you can change the location of the DATA directory, where the JEOL files are going to be placed. In case the user wants to merge different files that are supposed to be written with identical beam currents and different doses, set Geo stop to "YES". If this feature is not used one can usually just skip this menu and select "Continue". The computer will start crunching and, after it is done, will spit out huge files in a subdirectory called /DATA. If the

Geo stop is set, the data won't be written right into the DATA directory but into separate *.geo files, which have to be merged using the merge program.

- In case the above mentioned feature of merging different layers into one layer is used this has to be done at this point. From the command line, simply type "merge". The program asks for the layers to be merged. If all layers are supposed to be merged, type "*.geo". The program then asks for the shot ranks of each of the layers to be merged. This shot rank is assigned a different dose at runtime.

- Change to this directory and check where the field boundaries go through your layer by typing "xpreview layer1.j51 layer2.j51". Multiple layers can be previewed, in this case two layers are previewed.

- There are two very helpful programs on submicron.ll which help estimate the writing time for the e-beam. The first is called stats_51, a program that, when called as in "stats_51 layer1.j51" calculates the total area of boxes and length of lines to be exposed in the e-beam system. Those numbers should be written down for every layer. The second program is called calc and is invoked from the command line by typing "calc". In this program one can input the area of boxes and lengths of lines and receives an approximate time for the writing. The program demands additional information as to which lens is used (5th), which dose (e.g 400 μC/cm^2 and 3 nC/cm), and which current. The calculated times should be written down for every layer.

- If the field boundaries cut through your sensitive features you should go back and reconvert your file for a different selection of the origin. It might take a couple of iterations.

- After you are satisfied with the field boundaries, run the conversion program once more after typing "script". This will cause the whole output of the con-

version program to be piped into a file called typescript. After the conversion program is done exit the script and move typescript to the data directory under the name "logfile", or "script_date".

- Ftp all the data files, the logfile and the setup.cnv file to the nrl account on the NeXT machines. The ftp connection should be made binary (type "binary"). If you don't know the password of the nrl account, ask the system administrator. Again, if there doesn't exist a subdirectory with your name, create it.

- Remove all the files that are not necessary on the submicron.ll account. The DATA directory sometimes contains huge amounts of data and its contents should be removed if at all possible. If you want to keep anything in your directory for a while, like your cif file, please compress it.

- Log out of submicron and log into the nrl account on the NeXT machines. Go into your subdirectory.

- Run the program that converts the .j51 format into an equivalent VAX-readable format. NRL reads the data into their machine, which happens to be a VAX. The program can be started from your subdirectory by typing "../convert/layer > junk". The addition "> junk" is my creation because I wanted to suppress the screen output which dramatically slows down the conversion program. After all my conversion is done I delete the file. The resulting files should have an extension ".vax" and should have the same size as the corresponding ".j51" files.

- Please don't forget to immediately compress all the ".j51" files since they are not needed any longer. The ".vax" files should be compressed or deleted once the pattern is transferred to NRL.

- Send email to somebody at NRL so they can pick up the files and transfer them to their machine.

- print out a copy of the file "/NSL/Procedures/E-beam/KIC—>J51.ai". This table needs to be filled out with the layer specifications and desired beam currents so folks at NRL get an idea of what they should do with your files. Send or fax this sheet to NRL. Alternately, a document can be written and transferred to the nrl account as a PostScript file.

The following is a compilation of notes written down while Bill was showing us around the on the NRL Vax.

The following steps can be done only by persons who are authorized to use the account on NRL's Vax. The machine name is mpfvax.nrl.navy.mil (128.60.15.31), the account used for storing the data is called "mitcad". This machine runs VMS, so everything you have learned about UNIX is obsolete on this account. Translations of most common commands: "ls"→"dir"; "rm"→"del". VMS has a feature that prevents overwriting of old files by assigning a version number to each file.

- Ftp the .vax files to the NRL machine ("ftp mpfvax.nrl.navy.mil", user name MITCAD). Set ftp to binary. This process can take hours because the files are frequently very large and there is no common compression scheme between UNIX and VMS to the best of my knowledge. I have the strong suspicion that the file transfer is significantly faster if initiated from the remote machine.

- Log into mpfvax.nrl.navy.mil into the MITCAD account.

- Verify that the files you just sent were received by typing "DIR *.VAX".

- If there is a previous version of the file to be written, delete this file by typing something like "del NAME.vax:1"

- The .vax files are blocked off in 512 byte blocks. They need to be blocked off in blocks of 4096 bytes. This is accomplished by running the "F3" program. First check to see what the largest-number JEOL51*.DAT file is. To do this, type "dir jeol51*". For example, if the largest number is JEOL51080.DAT;1,

then the next data will be converted into JEOL51081.DAT;1. To initiate this conversion, type "RUN F3". This program will ask you for the input file. Type "<filename>.VAX". Then it will ask you for the output file. If the last unused JEOL51 file is JEOL51080.DAT, then type "JEOL51081.DAT". Repeat this for each file. For the next file, put the output in JEOL51082.DAT, and so on.

- To "fix" the directory, run the fixing program by typing "RUN J51DIRFIX".
- Verify that the new files have been updated by typing "TYPE JEOL51DIR-.DAT".
- To figure out the chip-size of each layer, run the program PATMNG. Give in "1" (for list-full) and "2" (for Disk No. :DUA1:) and answer the questions. This information was given to me by Kee Rhee on 2-16-94.

After all this is done the operator at NRL should be able to easily preview the pattern.

Appendix C

Electroplating

Power Supply

In order to electroplate x-ray masks, it is important to use a well-controlled and reliable plating setup. One of the essential parts of the setup is a plating current source. This appendix describes the design of the circuit.

There are two principal ways to electroplate a metal onto a substrate, which are referred to as galvanostatic and potentiostatic electroplating. During galvanostatic plating, the current density onto the substrate is held constant, while for potentiostatic plating, a voltage, e.g. the voltage between cathode and a saturated calomel electrode is held constant. This power supply was designed to feature both modes of operation.

Figure C-1 shows a schematic of the plating power supply. It was built for battery operation, but it can easily be converted for use with an outside power supply. The plating bath appears on the top left of the diagram, with pins 12 and 14 connected to the cathode and anode, respectively. The cathode is almost grounded, with less than $10\,\Omega$ to ground, due to the simple current divider into the current amplifier, which measures the plating current with 10 mV/mA amplification. The anode is driven by a 2N3904 transistor because the OP97 used to drive the transistor does not supply an output current above about 30 mA. The OP97 regulates either the saturated calomel

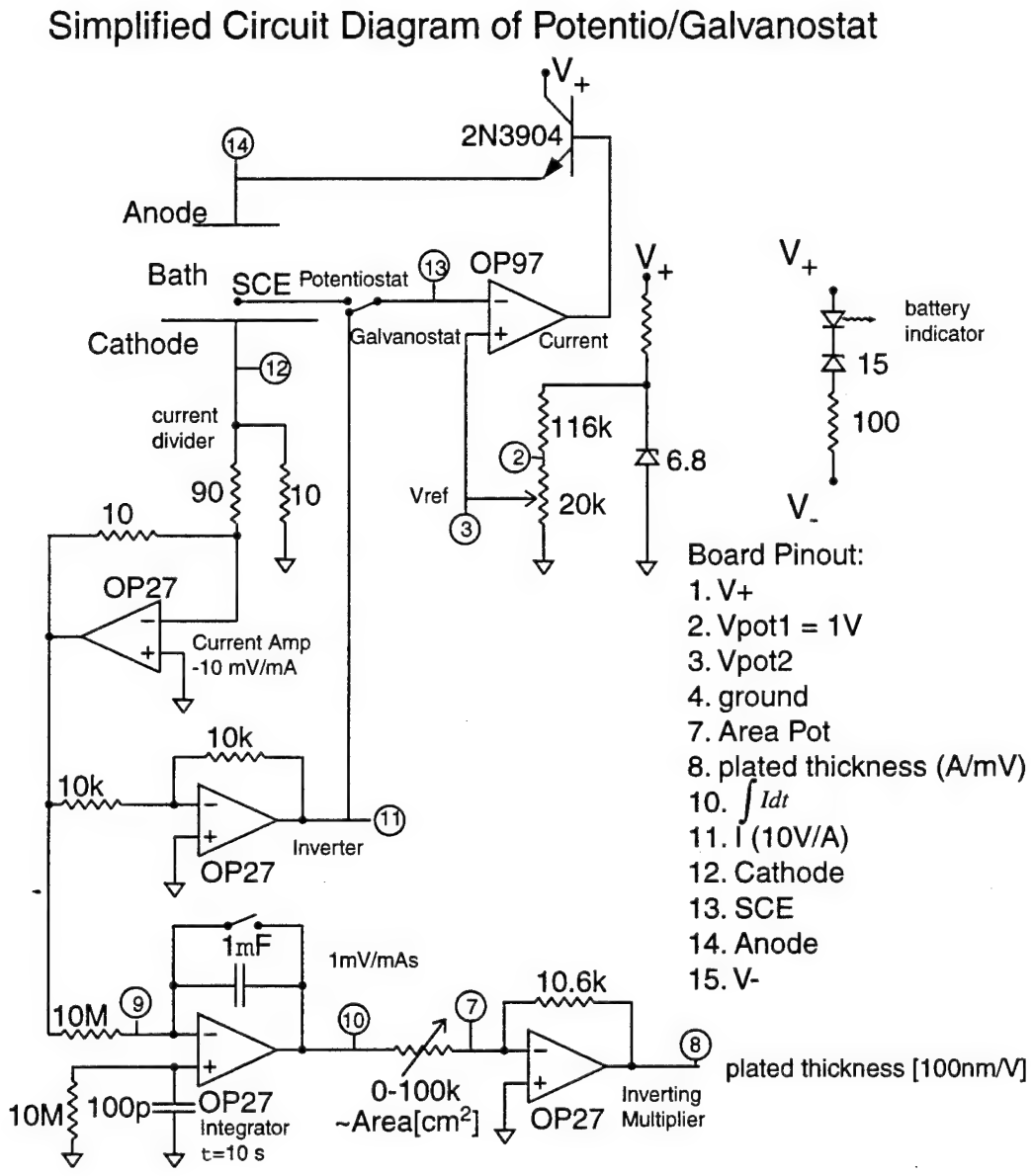


Figure C-1: Schematic circuit diagram of the plating power supply.

electrode (potentiostat), or the voltage measured with the current amplifier to be equal to the reference voltage which is set using a 6.8 Zener diode and a 10-turn potentiometer. The circuit on the bottom consists of an integrator, which integrates the current over time, $\int Idt$, and a successive amplifier. If the ten turn potentiometer between pins 7 and 10 is set to be proportional to the plated area and the plating efficiency is unity, the voltage on pin 8 gives the plated thickness, in Å/mV. The plating efficiency indicates how efficiently the ionic current in the solution transfers metal onto the sample. A plating efficiency of unity means that, for a unit charge injected into the solution at the cathode, one gold atom is plated out of the solution.

For trouble shooting, please use the schematic of Fig. C-1. All elements used in this circuit are described in basic circuit books [97,98].

There are several points that could be improved on this box:

- There seems to exist a high current latch-up state. It might be possible to avoid this state if the inputs of the OP97 are separately buffered.

- Change from battery operation to an external DC power supply. Those power supplies are commercially available, and it shouldn't take more than an hour to change the box accordingly, once such a power supply is obtained.

- It would be nice if the plating current and the plated thickness could be read out directly. It takes ordering of parts, in this case two miniature digital voltmeters, to complete the job.

Note: to use the power supply, one should turn the power supply on only if the sample is disconnected and the current setting potentiometer is zeroed. This will avoid above latch-up condition. Also, the circuit will latch if the current drawn is more than ~45 mA.

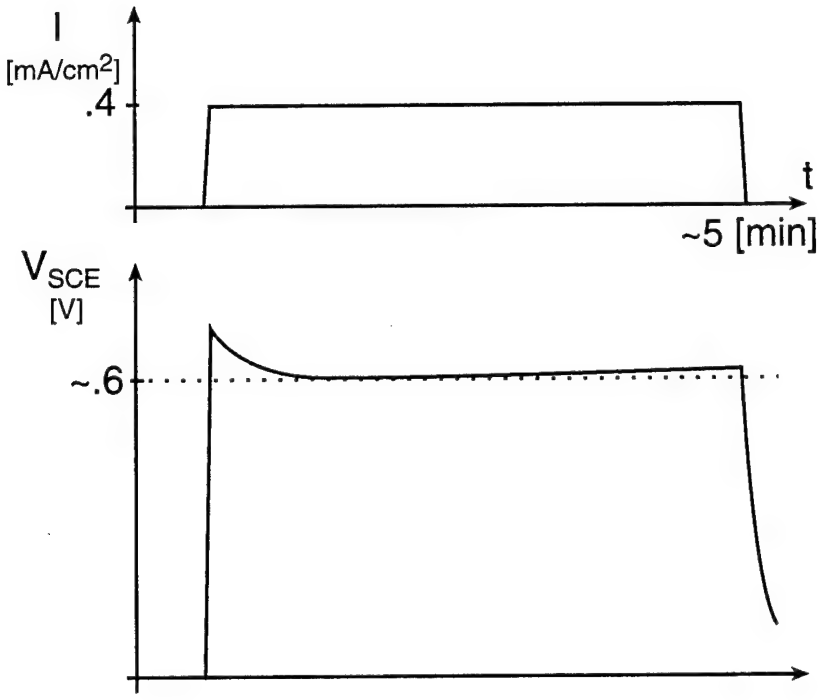


Figure C-2: Qualitative description of the SCE voltage as a function of time and plating current density.

SCE voltage during plating

The voltage of the saturated Calomel electrode should be monitored during the plating at all time using a strip chart recorder. This way it is easier to detect possible problems if the voltage deviates from previous, similar runs. A qualitative desciption of the saturated calomel electrode voltage as a function of time and plating current is given in Fig. C-2. After turning on the current, the voltage initially decays from the starting voltage, and starts rising a again after about two minutes. The rise in the end is probably due to depletion of the solution close to the cathode.

Design Considerations for Plating Fixture

A description of the plating apparatus used in plating masks was already given in section 2.5.1. Here are some comments about the design that should be incorporated in future fixtures and tanks.

It would be a good idea to construct the tank out of plexiglass or PVC, materials recommended by SEL-REX. Glass or pyrex beakers, as used in my experiments, are not recommended by the manufacturer, although it is not on the list of materials to avoid either. It might be a good idea to avoid these materials because of ionic impurities that might be present. A new tank should also provide space for the access of at least two sensors, the saturated Calomel electrode and a thermometer. I believe that it is a good idea to minimize the amount of plating fluid to reduce costs in case of a bath replacement as well as to reduce temperature stabilization times.

Notes on the bath

The following are the manufacturers recommendations for Operation of the bath in an abbreviated form:

- pH value has to be maintained between 8.5 and 9.0. This is, in my experience, rarely a problem, but it can be regulated by addition of sulfuric acid or sodium hydroxide.
- dull deposit $\rightarrow$ add 2 ml/l brightener.
- if that doesn't work $\rightarrow$ add conducting salts, $\sim$7g/l.
- if that doesn't work $\rightarrow$ more serious problem, possibly organic contamination.
- a sample should be periodically sent to be analyzed for gold content and conducting salts.

Appendix D

Fabrication Sequence

Following are the lithography steps used to make the devices described in this report. All steps are done with Microposit 1813 as resist, except for the gate metallization, where PMMA is used.

- Solvent clean the sample.
- Spin Microposit 1813 resist - I use 4.3 krpm for 30 sec $\sim 1\,\mu$m.
- Bake in oven at 90° C for 30 minutes.
- Perform alignment in OAI and expose for appropriate time, typically about 4 sec.
- Develop in Microposit 351 developer at room temperature.
- Perform mesa isolation of devices. I use 1000:20:6 H_2O:NH_4OH:H_2O_2 for 22 sec as the etch and typically get a 65 nm mesa (actually, sometimes I get up to 120 nm).
- Do exposure as above using ohmic contact mask level.
- UV ozone for 15 s.

- Rinse in cold DI water for 1 min.
- Rinse in Semico for 1 min. Semico is a buffered ammonium hydroxide solution that is supposed to strip any native oxide on the GaAs surface.
- Rinse in cold DI water for 1 min.
- Blow dry.
- E-beam evaporate ohmic contacts: Ni(2):Au(4):Ge(22):Au(44)Ni(10):Au(30).
- Lift off and anneal sample at 430°C for 15 s. in RTA.
- Solvent clean the sample.
- Spin KTI PMMA 950 K 4%, 5800 rpm for ~250 nm. I lift off up to 120 nm with this thickness.
- Bake PMMA at 180°C for 1 hour.
- Expose sample for a dose of approximately 600 J/cm^3, as determined by the x-ray database program, Appendix G.
- Develop sample for 5 seconds in standard developer and check development in alpha step. The developer is 2:3 MIBK:IPA and the development is performed using a squirt bottle. A total development time of about 30 to 60 seconds is desirable. Check development also in the microscope. The gas relief grating (see Appendix A) should be visible at this point, and in the areas of thicker polyimide the development should be a little reduced.
- Continue developing until the development is complete and add 10 seconds. Inspect again in optical microscope and check that the gas relief grating is not visible any more.

- UV ozone clean sample for 30 seconds to remove possible scum layer. Failure to do so might result in electrically unstable gates.

- Before evaporation of the gate metal, the native oxide on the surface of the semiconductor has to be removed. The sample is dipped for one minute into DI water, then one minute into Semico solution, then one minute again into DI water. As an alternative to Semico clean, a mixture of H_2O:NH_4OH can be used.

- Use thermal evaporator to evaporate 80 to 100 nm of gate metal. One can either use Aluminum or Gold/Palladium. Use of an electron beam evaporator might result in damage to the two dimensional electron gas. There are also a lot of successful device results made with e-beam evaporators. For electron beam evaporator, I usually use evaporation of 10 nm titanium and 90 nm gold.

- Perform liftoff by heating acetone on hotplate for 30 sec., putting sample in for 1 min. and then squirting sample with acetone from the plastic squirt bottle. Caution, heating acetone is like heating gasoline. For difficult liftoff, ultrasound may be used. Another solvent that can be used for liftoff is NMP, which is also a little less dangerous when heated. (Arvind's recipe, I believe he learned it at IBM: perform liftoff by soaking sample for 30 minutes in a mixture of 1:1 acetone:dichloro-methane.) The devices are now finished.

- Cleaving: put the piece down onto a paper towel and put a glass slide on each side. Put a third glass slide across at the position where you want to cleave the sample. It might be necessary to use a stereo microscope if you want to cleave between devices.

- Scribe the sample close to the bottom edge using a carbon scribe. It is not necessary to scribe along the whole length of the cleaving line. GaAs is very brittle and will easily cleave along the crystallographic planes.

- Position the sample on the glass slide with the cleaving line along the edge of the glass slide. Exert a force on the outermost sample edges. The sample should break without the exertion of much force if the lever arm is large enough. The smaller the sample, the more difficult it is to cleave.

- Take a 44-pin header and place a small amount of silver paint on it. Be careful not to put too much onto the header as the silver paint might cover also the front side of your sample. Press the sample down until most of the silver paint underneath the sample has moved outside.

- My experience with bonding devices with the Kulicke & Soffa bonder of the ICL is that the settings of three for time, force, and power is a good bet to start. Much higher settings present a physical danger to the device, lower settings won't complete the bond.

- Mount the sample onto the bonder chuck. Make sure you bond only the wires which have been grounded first to the chuck. The chuck and the person doing the bonding should both be connected to ground via a cable. The grounding is especially important for devices made with e-beam lithography since the metal tends to be thinner than for devices made with x-ray lithography.

- For the first bond, position the bonder wedge appropriately considering the height of your header and sample. Position the mouse over the first header pad and press the mouse button. Adjust the search button to place the wedge slightly above the pad. Then go to the sample pad and place the wedge slightly above the pad as well.

- For the remaining bonds you can use the z-stage lever with your right hand and position the sample with the mouse.

- All the cables in the measurement setup are designed as twisted pair cables starting with the pair 1-2. If at all possible try to bond source and drain on

successive bonds, following this convention. This will reduce noise pickup due to magnetic fields during measurements.

Appendix E

Diagnostic Pattern for E-Beam Lithography

The diagnostic pattern was developed to be dropped into existing patterns that are written with the electron beam on x-ray masks. It can also be used by itself as a tool to evaluate the performance of the electron beam. The diagnostic pattern was developed with the help of Scott Silverman and Scott Hector.

The layout of the whole pattern is shown in Fig. E-1. The pattern consists of six identical layers (indicated as different grey levels) which can be assigned separate doses or currents that the operator wants to get information about. The different layers are called D1–D6 in the layout program. The pattern is designed to fit within one field of 80 μm $\times$ 80 μm commonly used in the e-beam system at NRL. Under normal circumstances, it should be possible to write the whole pattern within about two minutes. The individual patterns are described in the following paragraphs:

Single pass lines and star-burst pattern: The single pass lines shown in Fig. E-2 and as "1" in Figure E-1 were designed to give information about linewidth of single pass lines both free standing and in proximity with other single pass lines. The right angle in the lines insures that any astigmatism will be detected as the difference in

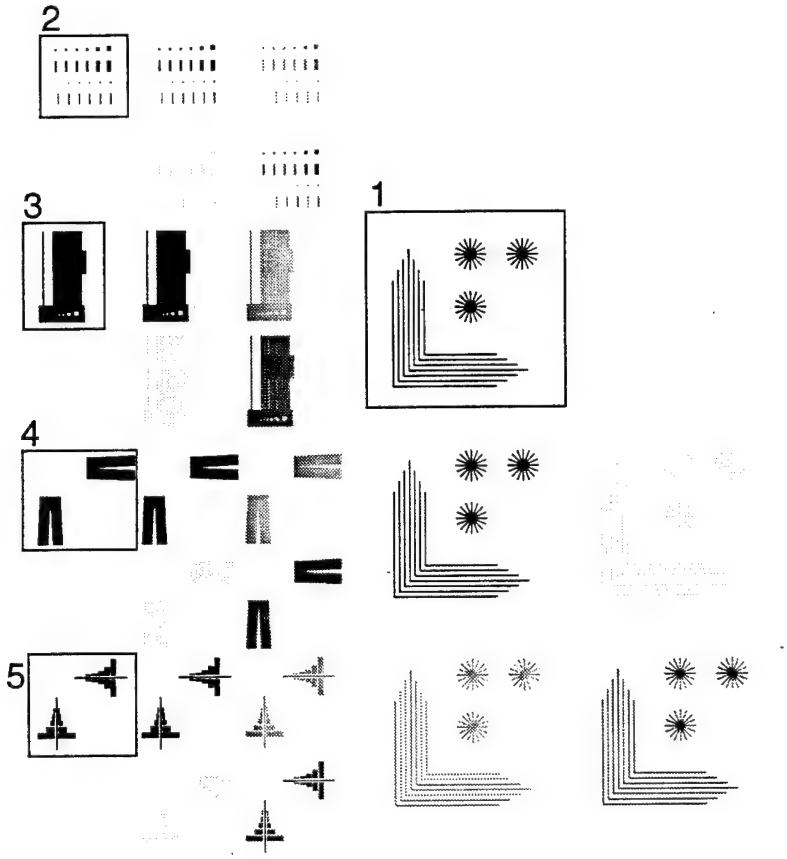


Figure E-1: Layout of the diagnostic pattern.

linewidth for x- and y-direction.

The so-called star-burst pattern was designed to write lines at angles of 30^o and 60^o in addition to lines in x- and y-direction. The intention was to show astigmatism by looking at the superposition of the lines. In reality, however, the lines at 30^o and 60^o receive slightly different doses than the x- and y-directed lines because of the digitization of the beam into squares. The dose for x- and y-directed lines tend to be higher. Since this pattern does not take very long to write, we believe the writing of the star-burst pattern is still worth the while.

Boxes and lines: This pattern, shown in Fig. E-3 and as "2" in Fig. E-1, was designed to print square boxes and lines using areal doses. The nominal linewidths

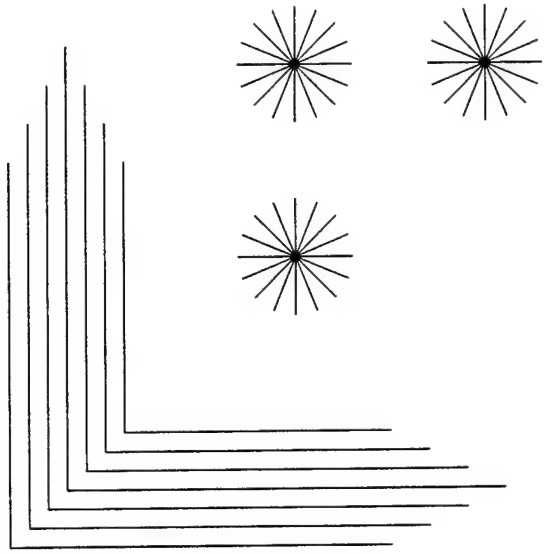


Figure E-2: Single pass lines and star-burst pattern in the diagnostic pattern.

and side lengths of the squares are 20 nm to 200 nm in 20 nm increments, plus 300 and 400 nm. The square boxes are of special significance, for example as contact holes in silicon processing.

Lines, grids, and inverted boxes: This pattern, shown in Fig. E-4 and as "3" in Fig. E-1 was designed to print boxes of inverted polarity and single pass lines adjacent to large areas, and to write grids and gratings of varying periodicity. The inverted polarity boxes have the same sizes as the boxes in pattern "2", but are of the opposite polarity.

Inverted Line: This pattern, shown in Fig. E-5 and as "4" in Fig. E-1 was designed to check the printing of an inverted line in the presence of large written areas surrounding the line. To check for stigmation, there is one pattern in x-direction and one in y-direction. The linewidth ranges from 20 nm to 400 nm in 20 nm increments. The first and last segments are 500 nm-long, all other segments are 200 nm-long. The width of the boxes on the sides of the line is 1 μm.

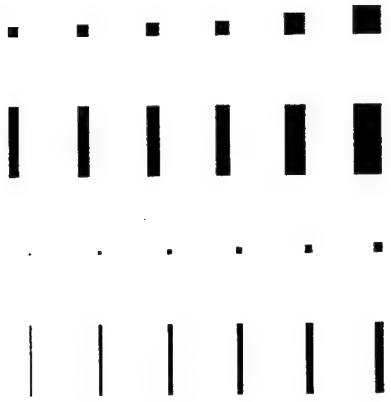


Figure E-3: Boxes and lines in the diagnostic pattern.

Effect of proximity effect on a single line: This pattern, shown in Fig. E-6 and as "5" in Fig. E-1 tests the proximity effect of adjacent boxes on a single pass line. The space between the boxes is held constant at 400 nm, the size of the boxes is varied from 100 nm, 200 nm, 400 nm, 800 nm up to 1.6 μm. To check for stigmation, there is one pattern in x-direction and one in y-direction.

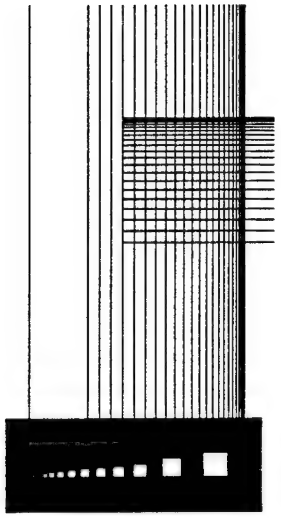


Figure E-4: Grid test pattern and inverted boxes in the diagnostic pattern.

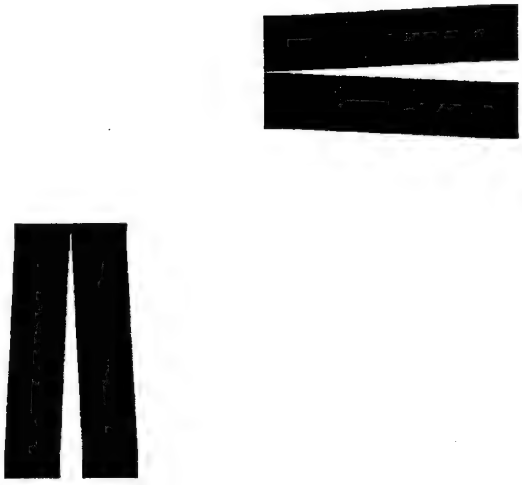


Figure E-5: Line of inverted polarity in the diagnostic pattern.

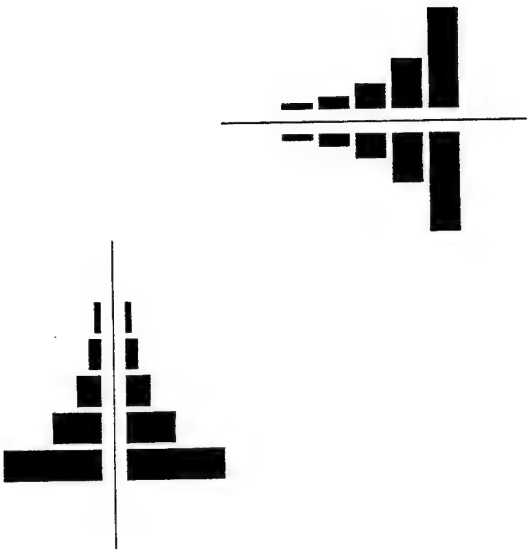


Figure E-6: Proximity effect test structure in the diagnostic pattern.

Appendix F

Kic Layout Tools

This chapter describes the layout tools that were developed in the course of the report research. The layout program used for design of devices is KIC. KIC is a relatively primitive editor which produces a file which consists of a short header, layer names for each new layer and a list of boxes, wires and polygons.

The command for a typical box is:

```
B 390000 5000 605000 390000;
```

This line describes a box of width 3900 λ and height 50 λ and center coordinates of x:6050 λ and y:3900 λ.

The command for a typical single pass line is:

```
W 0 0 6400000 30000000 6400000;
```

This line describes a line of zero width (single pass) and the x- and y-coordinates of the beginning and end points of the line.

Moirè Alignment Mark

For high accuracy alignment of x-ray masks, the alignment marks are of great importance. This section describes the code used to generate alignment marks as described

in ref. [99]. The alignment mark consists of gratings of two different periodicities which are perpendicular to two others. Such an alignment mark is shown in Fig. F-1. When aligning a mask to a substrate, a pair of two horizontal and vertical fringe-patterns each appear. The fringes of each pair move in opposite directions when the substrate is moved with respect to the mask.

The program to produce the alignment mark is listed below:

```
#include <stdio.h>
#include <math.h>
 /*this program creates an alignment mark for fine alignment of x-ray */
 /*masks . It expects an input file containing the periodicity of the */
 /*gratings. The input file should contain "period1,period2,size" of */
 /*the alignment mark. It produces triangular gratings with aspect */
 /*ratio 1:1. Lambda is going to be 10 nm and the gratings are to be a */
 /*triangles of base 80 um and 40 um height.  The horizontal grating */
 /*is going to be of period 1, the vertical of period2. I am not quite */
 /*sure if I have to account for the rounding error occuring from the */
 /*conversion from float to integer for nyc1 and nxc2 */
 /*written by Martin Burkhardt 1-94*/

main()
{
  float xc1,w1,yc1,h1; /*width, height, and center coordinates of boxes*/
  float xc2,w2,yc2,h2;
  int nyc1,nh1,nxc1,nw1;
  int nyc2,nh2,nxc2,nw2;
  int i,no1,no2,nwidth;
  int offset;
  double period1,period2,width;
  FILE *fp, *input, *fopen();
```

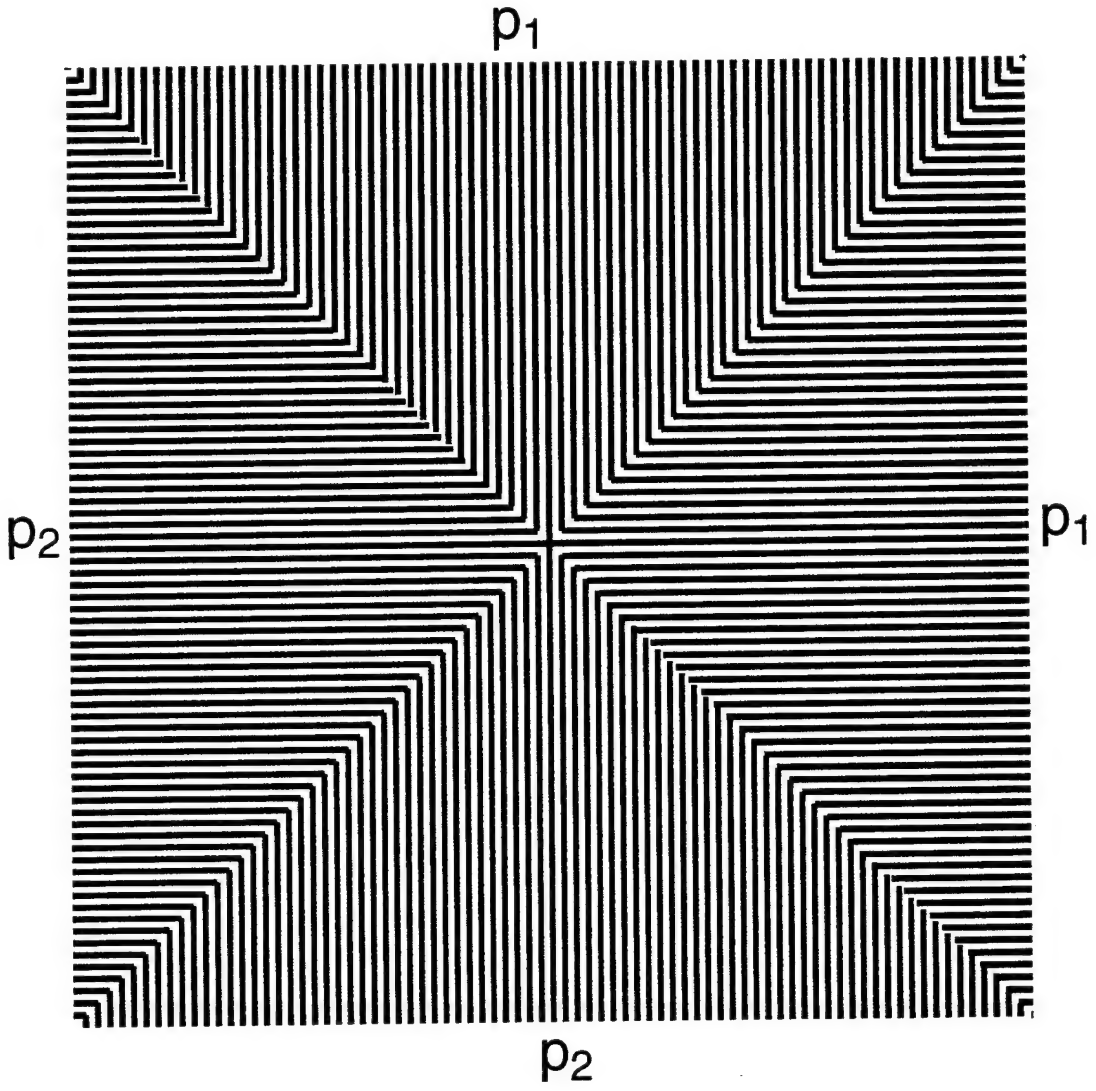


Figure F-1: Example of a Moirè alignment mark. Written with an e-beam, the outer dimension of the alignment mark would be 80 μm. The horizontal periodicity is 1 μm, and the vertical is 1.05 μm in this case.

```
fp = fopen("align.kic","w");
input = fopen("input.dat","r");

/*the following input is expected from the input file "input.dat"*/

fscanf(input,"%lf %lf %lf",&period1,&period2,&width);
no1 = width/(2*period1);
no2 = width/(2*period2);
period1 = period1 * 10000;  /*convert microns into lambda*/
period2 = period2 * 10000;
width = width * 10000;
nwidth = width; /*integer conversion*/
offset = nwidth / 2;

fprintf(fp,"(Symbol align.kic);\n");
fprintf(fp,"9 align.kic;\n");
fprintf(fp,"DS 0 1 1;\n");
fprintf(fp,"L HRA;\n");

/*central lines*/

xc1 = width/4; /*x-coordinate of center 1*/
xc2 = 0; /*x-coordinate of center 2*/
nxc1 = xc1;
nxc2 = xc2;
yc1 = 0; /*y-coordinate of center 1*/
yc2 = width/4; /*y-coordinate of center 2*/
nyc1 = yc1; /*might have to add half a unit to account for rounding*/
nyc2 = yc2;
```

```
h1 = period1 / 2; /*height of box 1*/
h2 = width / 2; /*height of box 2*/
nh1 = h1;
nh2 = h2;
w1 = width / 2; /*width of box 1*/
w2 = period2 / 2; /*width of box 2*/
nw1 = w1;
nw2 = w2;
fprintf(fp,"B %d %d %d %d;\n",nw1,nh1,nxc1+offset,nyc1+offset);
fprintf(fp,"B %d %d %d %d;\n",nw1,nw2,-nxc1+offset,nyc1+offset);
fprintf(fp,"B %d %d %d %d;\n",nw2,nh2,nxc2+offset,nyc2+offset);
fprintf(fp,"B %d %d %d %d;\n",nh1,nh2,nxc2+offset,-nyc2+offset);

/*all other lines*/

for(i=1;i<=no1;i++){
  xc1 = width/4 + i * period1 / 2; /*x-coordinate of center 1*/
  yc1 = i * period1;  /*y-coordinate of center 1*/
  nxc1 = xc1;
  nyc1 = yc1; /*might have to add half a unit to account for rounding*/
  h1 = period1 / 2;  /*height of box 1*/
  w1 = width / 2 - i * period1; /*width of box 1*/
  nh1 = h1;
  nw1 = w1;
  fprintf(fp,"B %d %d %d %d;\n",nw1,nh1,nxc1+offset,nyc1+offset);
  fprintf(fp,"B %d %d %d %d;\n",nw1,nh1,nxc1+offset,-nyc1+offset);
  fprintf(fp,"B %d %d %d %d;\n",nh1,nw1,nyc1+offset,-nxc1+offset);
  fprintf(fp,"B %d %d %d %d;\n",nh1,nw1,-nyc1+offset,-nxc1+offset);
}
for(i=1;i<=no2;i++){
```

```
    yc2 = width/4 + i * period2 / 2; /*x-coordinate of center 2*/
    xc2 = i * period2;  /*x-coordinate of center 2*/
    nyc2 = yc2;
    nxc2 = xc2;
    w2 = period2 / 2;  /*height of box 2*/
    h2 = width / 2 - i * period2; /*height of box 2*/
    nw2 = w2;
    nh2 = h2;
    fprintf(fp,"B %d %d %d %d;\n",nw2,nh2,nxc2+offset,nyc2+offset);
    fprintf(fp,"B %d %d %d %d;\n",nw2,nh2,-nxc2+offset,nyc2+offset);
    fprintf(fp,"B %d %d %d %d;\n",nh2,nw2,-nyc2+offset,nxc2+offset);
    fprintf(fp,"B %d %d %d %d;\n",nh2,nw2,-nyc2+offset,-nxc2+offset);
  }
  fprintf(fp,"DF;\n");
  fprintf(fp,"E\n");
}
}
```

Appendix G

X-Ray Lithography Run-sheet

This appendix describes the X-ray run-sheet developed for the electron bombardment x-ray sources in the Nanostructures Laboratory at MIT. The run-sheet uses a database file that runs off a NeXT computer in the lab. An example of the run-sheet is shown in Fig. G-1. The advantage of running the run-sheet off a network computer is that all the parameters for the run can be accessed from remote locations. Another advantage is that the parameters given for the run can be converted into doses with the use of corresponding formulas, and development times can be estimated. In this database, the dose incident and dissipated in the resist is calculated, the development time for the samples is recorded, and values for diffraction coefficient α and penumbra coefficient β are calculated for a given linewidth, gap, and source size. Furthermore, it is an easy task to summarize the data for several months of operation and to statistically evaluate the parameters.

The database was created with the help of Mark Schattenburg and Scott Hector.

Dose incident in the resist

The incident power P_b of the x-rays on the nitride membrane is given by

$$P_b = \frac{P_e B}{d^2}$$

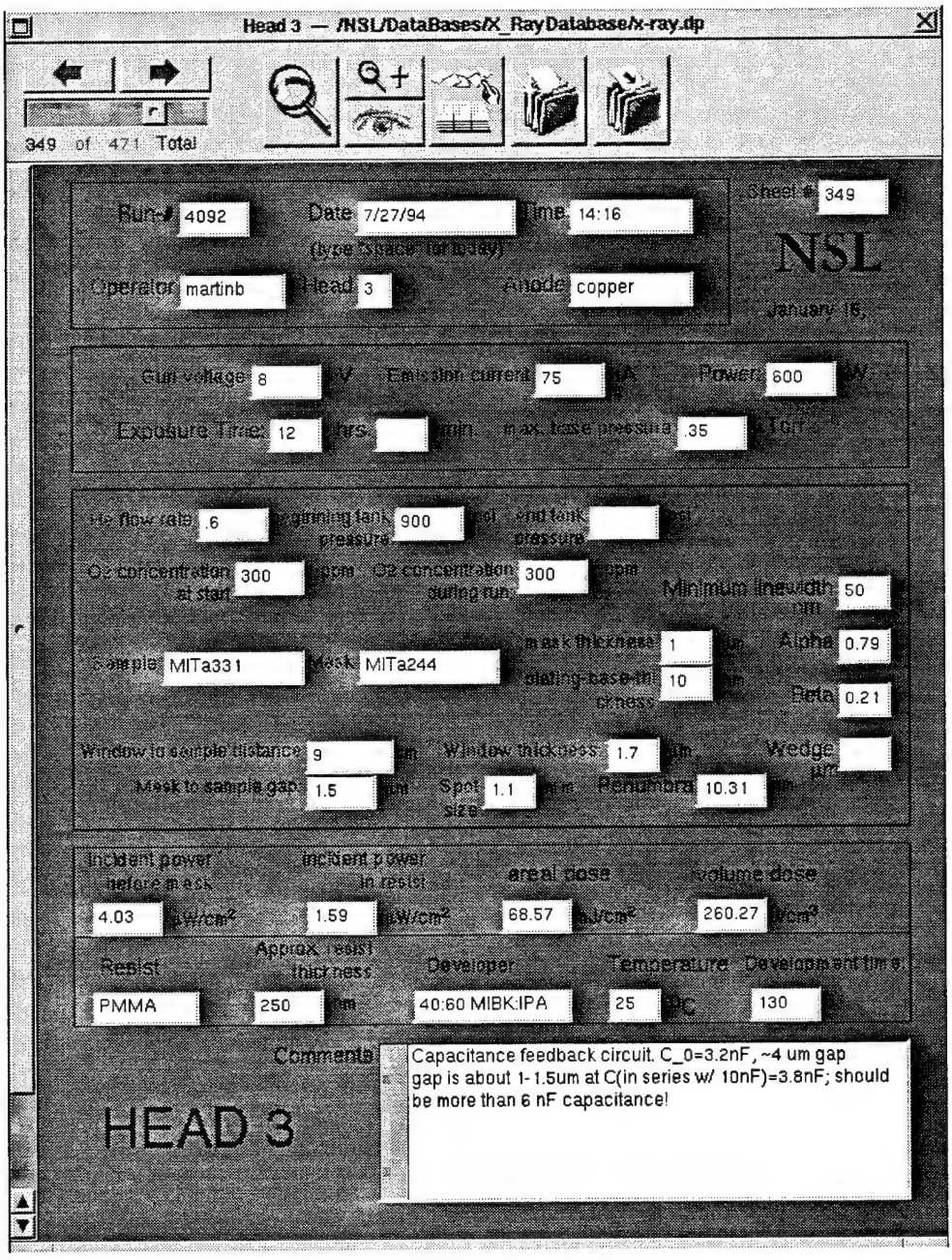


Figure G-1: Sample of a run-sheet for an exposure in helium environment.

which states that the incident x-ray power is the product of the electrical power P_e times the brightness of the source divided by the square of the distance. The brightness is given in units of W/SrW. The values for the brightness are taken from Henke [100] for anode voltages of 8 kV for copper and 6 kV for graphite, which are the voltages of maximum brightness, and given in tab. G.1.

Line Source	Brightness[ph/SrWs]	Brightness[W/SrW]
Cu-L	4.8×10^{10}	$7.15\times10^{-}6$
C-K	1.83×10^{11}	$8.14\times10^{-}6$

Table G.1: Brightness of X-ray Sources.

To calculate the areal dose, the power is multiplied by the total time of the exposure. For the volume dose, one has to calculate the absorption in the PMMA. The mass absorption coefficient in PMMA for Cu-L x-rays is m=3324.9 cm^2/g, the density of PMMA is $r = 1.2$ g/cm^3. This results in a coefficient $rm = 3990$ cm^{-1}. We calculate the areal dose at the resist-substrate interface and the volume dose is the difference between the two areal doses divided by the resist thickness. Here we assume that for a Taylor expansion of the exponential intensity decay, $e^{\alpha x} = 1+\alpha x+\alpha x/2+...$ we can neglect all but the constant and linear term.

Calculation of attenuation in Au, Cr, SiN, Polyimide, and Helium Atmosphere

The intensity of the x-rays is proportional to $I \propto e^{-\sigma t}$, where σ is the absorption coefficient and t is the thickness of the resist. We can write for the output power P_o, after traveling a distance t:

$$P_o = P_i e^{-\sigma t} \quad ,$$

where P_i is the input power. Attenuations are usually given for different materials in dB per micron for a specific wavelength. [1] The attenuation coefficients for the Cu-L x-rays, used in our lab, are given in tab. G.2. For completeness, the attenuation

Material	A[dB/μm]	σ[1/m]
Silicon Nitride	3.4517	7.948×10^5
Polyimide	1.9897	4.581×10^5
Au	49.694	1.144×10^7
Cr	27.248	6.274×10^6

Table G.2: Attenuation coefficients for Cu-L line.

coefficients for the C-K x-rays are given in tab. G.3.

Material	A[dB/μm]	σ[1/m]
Silicon Nitride	30.002	6.908×10^6
Polyimide	1.7384	4.003×10^5
Au	119.46	2.751×10^7
Cr	37.020	8.524×10^6

Table G.3: Attenuation coefficients for C-K line.

For x-ray exposures in a helium ambient, the attenuation in helium needs to be calculated. The transmission, T, is given by

$$T = e^{-\mu\rho t} \quad ,$$

[1] The conversion can be done using the definition for dB

$$A[dB] = -10 \times \log(\frac{P_o}{P_i}) \quad ,$$

that is 10 times the logarithm of output to input power. Since we have dB per micron given, we can equate above equations:

$$\sigma = -10^6 \ln(10^{-\frac{A[dB]}{10}})[\frac{1}{m}]$$

where ρ is the density of the gas. Since the chemical composition of air is

$$N_{79}O_{21}He_{(\frac{21}{\%O_2}-1)\times 50} \ ,$$

we can calculate for its density, ρ,

$$\rho = (.18 + \frac{\%O_2}{21}) \times 10^{-3} g/cm^3 \ .$$

The magnitude of μ can be determined from the formula:

$$\mu = \frac{\sum_A \mu_A S_A W_A}{\sum_A S_A W_A}$$

where S is the percentage of each gas in room air, and W_A is the atomic weight of each gas. The values for μ are tabulated (in cm^2/g) for Cu-L x-rays (λ =13.3 Å), and are given in tab. G.4. For this particular case, we can calculate the magnitude of μ:

Element	μ
N	3.9191×10^3
O	5.4575×10^3
He	7.1734×10^1

Table G.4: Attenuation coefficients for Cu-L line in gases.

$$\mu = \frac{79 \times \mu_{N_2} \times 14 + 21 \times \mu_{O_2} \times 16 + 50(\frac{21}{\%O_2} - 1) \times \mu_{He} \times 4}{79 \times 14 + 21 \times 16 + 50(\frac{21}{\%O_2} - 1) \times 4} \ .$$

It should be noted, however, that the doses calculated this way are not completely consistent among the three different x-ray sources in our lab. For head 3, which features a vacuum window and He-atmosphere, the doses are typically lower for the same development time than doses in head 1. This draws the absolute dose information calculated in this database into question. The need to develop an absolute dose

measurement system must be emphasized. It is still possible to compare exposures done on the same x-ray source.

Two additional parameters, α and β, were introduced into the database [18,101, 20]. α is a measure of diffraction, as defined in eq. (2.1):

$$G = \alpha \frac{W^2}{\lambda}$$

G is the maximum allowable gap, W is the minimum linewidth, and λ is the x-ray wavelength. β is a measure of penumbra δ, defined by the formula

$$\delta = \beta W \ ,$$

which states that β is the ratio of the penumbra to the minimum linewidth. The penumbra can be calculated if the source extent, the source-to-sample distance, and the gap are known.

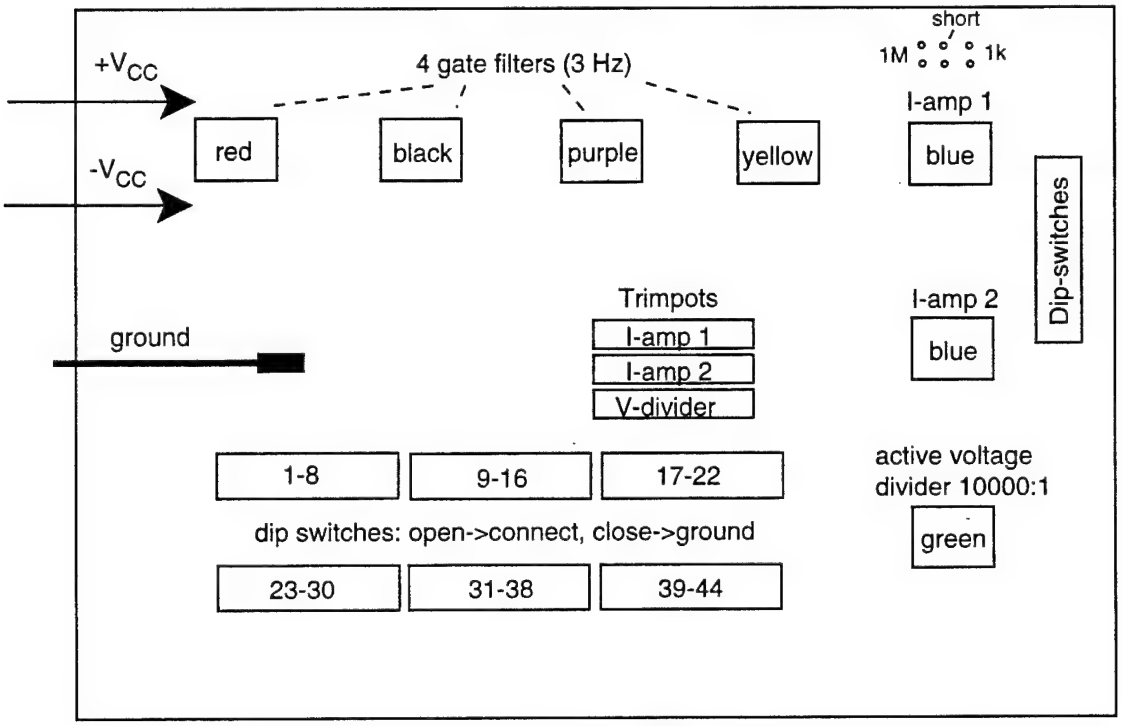
Appendix H

Electrical Measurements Box

In order to perform low-temperature measurements at low noise levels, it is necessary to design an interface between the probe and the measurement equipment. The basic design of this interface was done with the help of Arvind Kumar and Nathan Belk. A good reference for low noise measurement systems was written by Ott [102].

Board layout: The layout of the board is shown in Fig. H-1. The interface box consists of four gate filters, an active voltage divider and two current amplifiers. To keep the layout clear, most of the circuitry was concentrated on the backside of the board. Only the operational amplifiers (opamps), trimpots, connectors, and dip-switches were put onto the top of the board to facilitate easy servicing and switching. All wires that need to be connected to the board are brought onto the board through connectors so that it is relatively easy to take out and service the board. All circuits are assembled on one circuit board and tied to a single ground point.

The four low-pass filters for the gates have a cut-off frequency of 3 Hz and have color coded wires and usually use OP27 opamps. The colors used for the wires of the corresponding sub-circuits are indicated on Fig. H-1. A circuit diagram of the gate filters is shown in Fig. H-2. The active voltage divider, whose circuit diagram is shown in Fig. H-3, reduces the voltage by a factor of 10,000 and filters the voltage



2 9V transistor batteries
lifetime ~ 30 hrs.

Dip switch assignment:
1-1e6 V/A for I-amp 1 when closed
2-1e8 V/A for I-amp 1 when closed
3- NC
4- series-C for I-amp 1 when open
5-1e6 V/A for I-amp 2 when closed
6-1e8 V/A for I-amp 2 when closed
7- NC
8- series-C for I-amp 2 when open

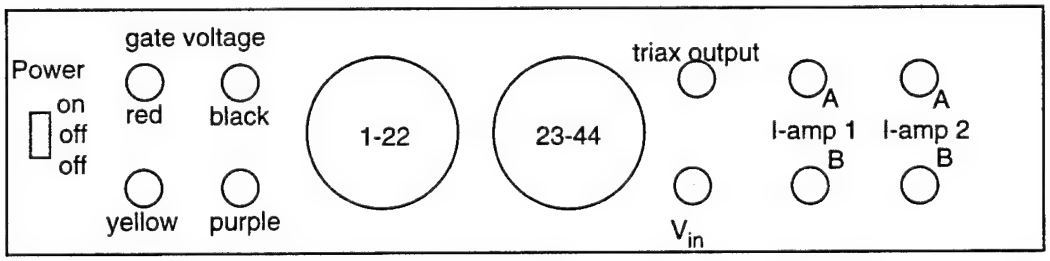


Figure H-1: Layout of interface box.

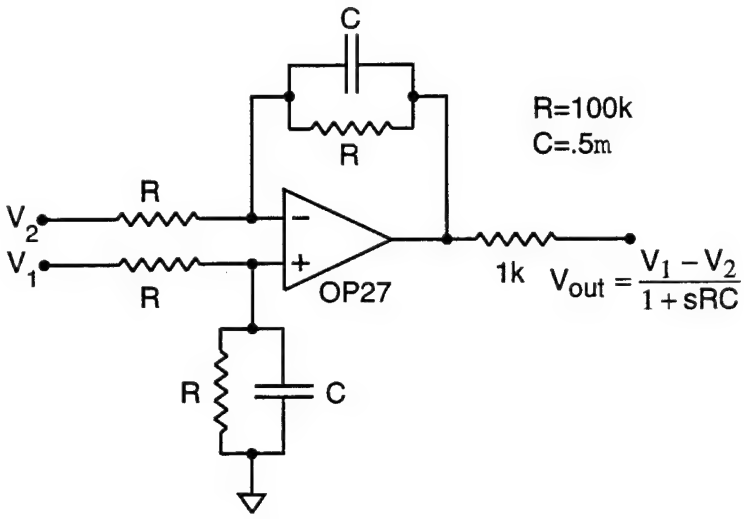


Figure H-2: Circuit diagram of gate filter.

at the same time with a 160 Hz low-pass filter. The voltage divider uses OP27 as well; any offset voltage can be eliminated using a trimpot. The current amplifiers, shown in Fig. H-4, have a gain of either 10^6 or 10^8V/A, which can be selected using the dip switches in the right top corner of the board. Another dip switch connects a 10 μF capacitor in series with the input of the current amplifier. The capacitor is an AC-short for high device impedances, which is the case for transistors below threshold. The board is powered with two 9 V transistor batteries, which have a hold time of about 36 hours under a load of all six opamps as shown in figure H-5. The 44 leads of the two cables are connected to the dip switches on the bottom of the board. A closed dip switch connects the corresponding lead to ground, an open dip switch floats the lead that can now be set to any potential. On the top right of the box are some connectors, which can be used to calibrate the voltage divider and current amplifiers. The two middle contacts are simply shorted so that one can directly measure the applied voltage by connecting the green wire (output of voltage divider) to one of the black wires. In between the right two contacts there is a 1 kΩ resistor and between the left two contacts a 1 MΩ resistor, which can be used to calibrate the current amps by connecting the green wire (voltage divider) and the

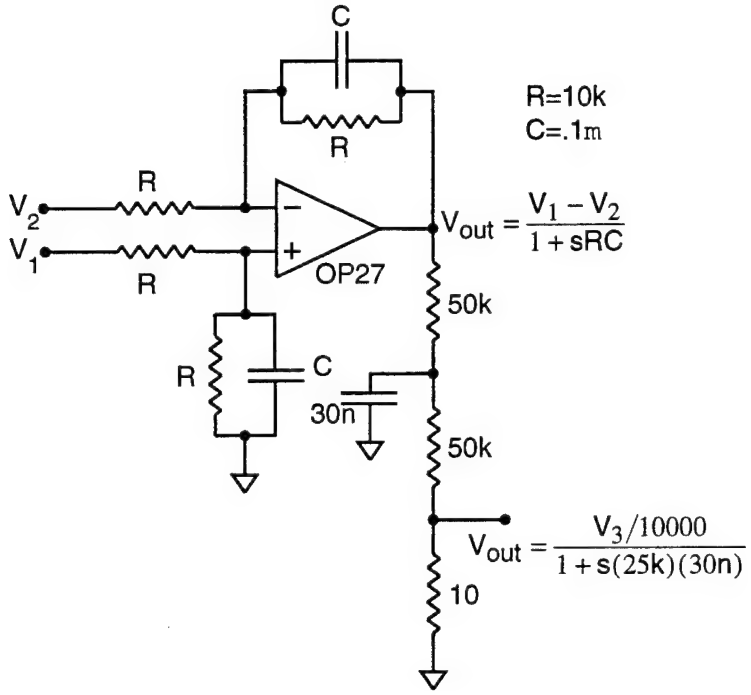


Figure H-3: Circuit diagram of active voltage divider.

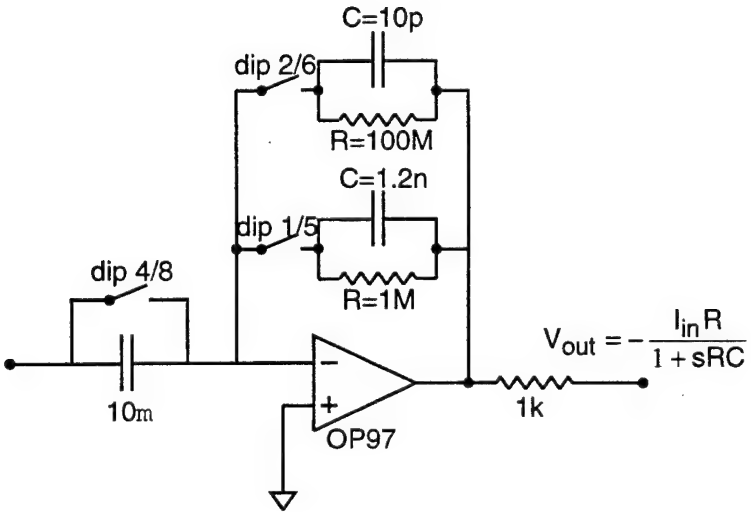


Figure H-4: Circuit diagram of current amplifier.

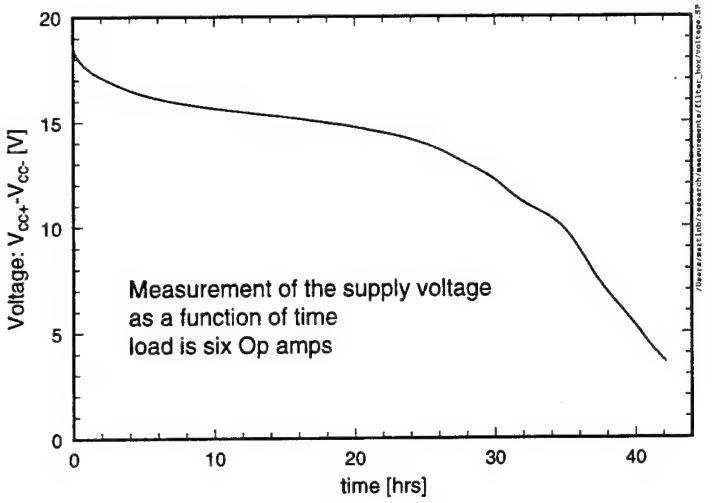


Figure H-5: Supply voltage versus time.

blue wire (current amp) and reading out the current value.

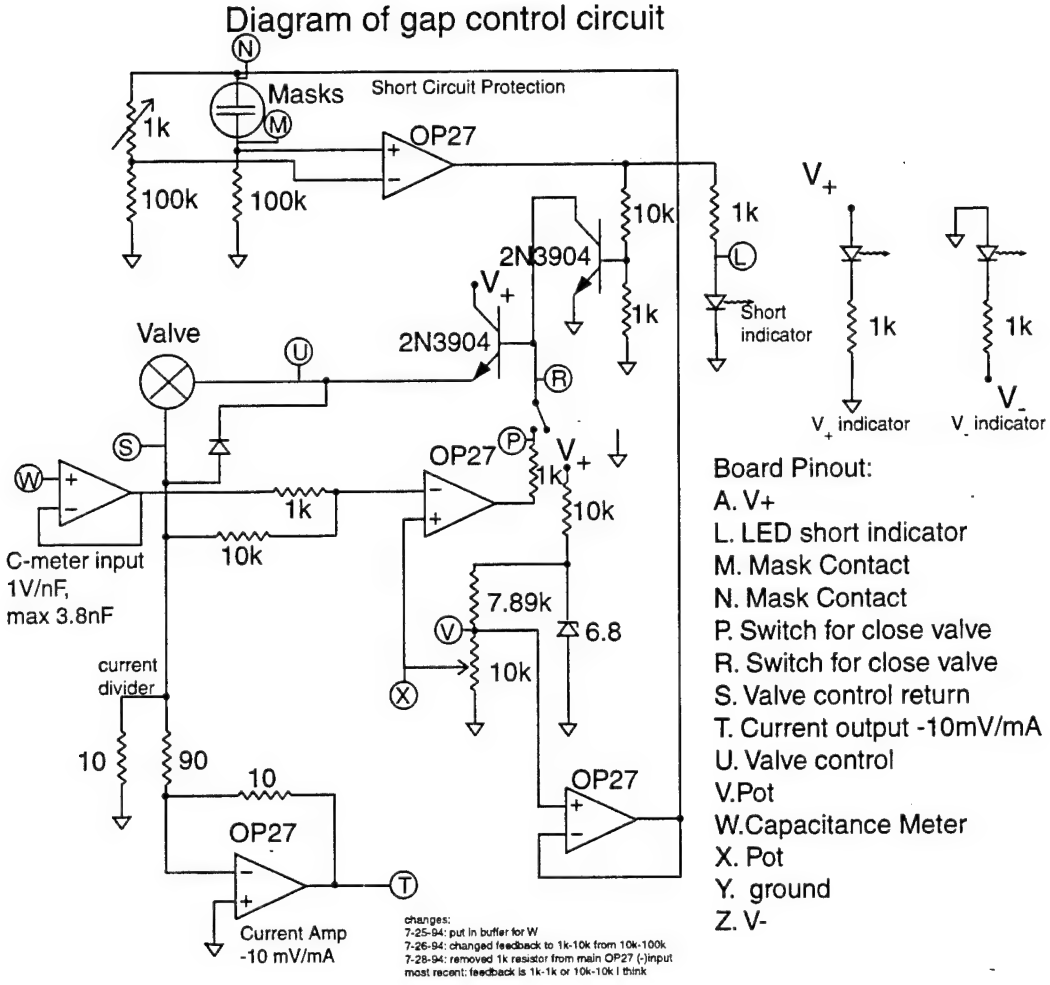
Appendix I

Gap Control Circuit

Circuit description

The circuit that enables one to control the gap between the mother and daughter masks during mask replication is shown in Fig. I-1. It is a straightforward regulating circuit using operational amplifiers to keep the capacitance between the masks constant. For simplicity and availability OP27s were used as operational amplifiers (opamps), but more primitive opamps such as the OP07 could be used as well. A Boonton 72B capacitance meter is used to measure the capacitance and its analog output is supplied to the circuit on pin W. This voltage, weighted by the feedback resistors, is compared to the reference voltage which is set by the operator, using a variable voltage divider and a Zener diode, at the central operational amplifier. This opamp drives an *npn*-transistor which regulates the MKS 248A valve to keep the capacitance constant. The transistor can be switched off manually or automatically in case a short between the masks occurs. The short is detected using the opamp in the top of the diagram which acts as a switch in a conventional Wheatstone bridge configuration. The opamp on the bottom measures the current going through the valve for diagnostic purposes.

Not shown in this figure is a separate capacitor which can be switched in parallel with the masks. This has the purpose of extending the measurement range of the



There is a capacitance in series with the mask to extend the range of the C-meter. This capacitance is about 9.44 nF, nominal value is 10 nF.

Valve specs: current controlled, completely open at less than100 mA, should be protected to less than 140 mA. There is a hysteresis, and the threshold is around 30 to 50 mA, depending on the setting on the bottom. It takes about 10 to 15 mA from turn on to completely open. The internal resistance of the coil is 100 Ohms, so that the current is automatically limited to 137 mA in this case because of V+ = 15 V. I=15V/109Ohms = 137 mA!

Figure I-1: Diagram of gap-control circuit.

capacitance meter above 3.8 nF. The series capacitance has a nominal value of 10 nF, and the capacitance can be extracted from the displayed capacitance using $C = 1/(1/C_{disp} - 1/10\text{ nF})$. This way, the mask capacitance can be extended to about 6.1 nF.

Operation

First, the daughter mask should be checked for particles and placed, membrane side up, onto the o-ring on the metal base plate, after being connected to one of the contact leads. Then the mother mask should be checked for particles, connected to the other contact lead (polarity doesn't matter) and placed membrane side down onto the daughter mask. The capacitance meter should be connected to a strip chart recorder and monitored during the entire run. The gap can be monitored easily with green light, as described in 2.3.3. After about a minute, the masks should be at a gap determined by the mesa warp or by aluminum studs, typically about 3-5 μm. To verify that the leads make contact to both masks, the capacitance should read approximately 2 nF. In case there is a short, the LED on the box will light and the capacitance meter will be off scale. In this situation, the mask should first be rotated by a small amount to see if the masks are still shorted together. If this does not work, either particles have to be removed, or a different daughter mask has to be used. The potentiometer in the front of the box should now be dialed until the valve starts to open up, which can be determined by observing the interference fringes in green light. Alternatively, the valve current can be monitored using the second channel of the strip chart recorder. The capacitance can be increased up to 3.8 nF without the series capacitor mentioned in the previous paragraph. The gap can easily be set to about 1 μm. Below a gap of 1 μm there is a possibility that the masks will come into intimate contact, which is not a desirable state. During exposure, resist on the daughtermask outgases and creates bubbles and large gaps associated with it. Contact exposures typically don't turn out successful unless there is some provision for the

gas to escape. Furthermore, if the masks are in intimate contact, they typically stay that way until forced apart, endangering both the mother and the daughter mask. A state in which the masks are in intimate contact appears to be energetically favorable over a state in which the masks are at a small gap. The reasons for this effect have not been fully explored. To get the masks out of intimate contact, I typically use thin wires to exert a small lever on opposite sides of the masks to force them apart.

Appendix J

Calculation for Double Dot System

This appendix contains a calculation of the charging energy using the method used by Ruzin et al.[89]. In order to keep calculations consistent with section 3.5, the indices are adjusted to that convention. The two dots and the charges in the dot are shown schematically in Fig. J-1. The charging energy U depends on the charges q_i on the capacitors C_i in the following way:

$$U(q_{ag}, q_{bg}, q_{al}, q_{bl}, q_{ab}) = \frac{q_{ag}^2}{2C_{ag}} + \frac{q_{bg}}{2C_{bg}} + \frac{q_{al}^2 + q_{bl}^2}{2C_l} + \frac{q_{ab}^2}{2C_{ab}} - V_g(q_{ag} + q_{bg}) \ . \qquad \text{(J.1)}$$

C_{ag} and C_{bg} are the capacitances between the gate and dot a and b, C_l is the dot to lead capacitance, and C_{ab} is the dot to dot capacitance. Because the charge in the dots is discrete, charge balance must be fulfilled[1],

$$\begin{aligned} q_{al} + q_{ab} - q_{ag} &= -N_a e \\ q_{bl} - q_{bg} - q_{ab} &= -N_b e \ . \end{aligned} \qquad \text{(J.2)}$$

[1]Ruzin [89] uses $q_{bl} + q_{bg} - q_{ab} = -N_a e$, which is incorrect with the chosen sign convention.

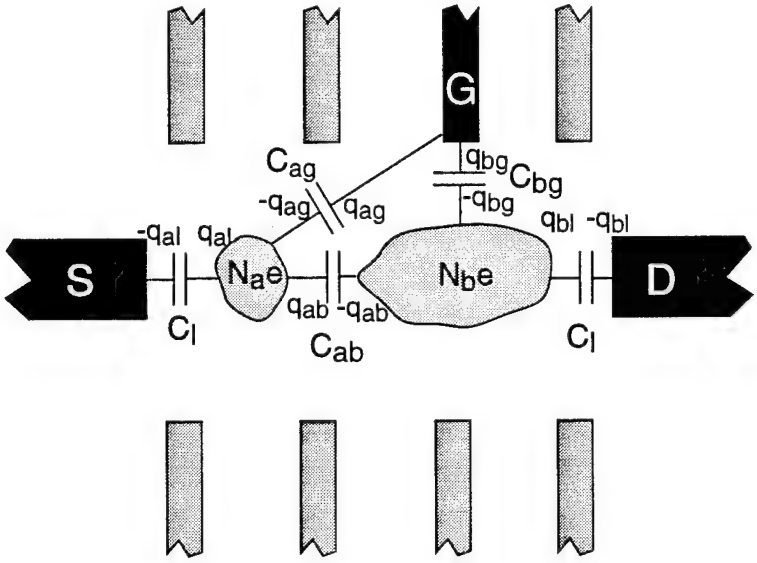


Figure J-1: Schematic of the double dot device, depicting the important capacitances in the double dot system.

For every state, the charging energy should be at a minimum, so that we get

$$\partial U/\partial q_i = 0, \quad i = al, bl, ab \ . \tag{J.3}$$

This leaves us with five equations for five unknowns and we have to solve the following matrix equation

$$\begin{pmatrix} 1 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & \frac{1}{C_l}+\frac{1}{C_{ag}} & 0 & \frac{1}{C_{ag}} \\ 0 & 0 & 0 & \frac{1}{C_l}+\frac{1}{C_{bg}} & -\frac{1}{C_{bg}} \\ 0 & 0 & \frac{1}{C_{ag}} & -\frac{1}{C_{bg}} & \frac{1}{C_{ab}}+\frac{1}{C_{ag}}+\frac{1}{C_{bg}} \end{pmatrix} \begin{pmatrix} q_{ag} \\ q_{bg} \\ q_{al} \\ q_{bl} \\ q_{ab} \end{pmatrix} = \begin{pmatrix} eN_a \\ eN_b \\ V_g - \frac{eN_a}{C_{ag}} \\ V_g - \frac{eN_b}{C_{bg}} \\ \frac{eN_a}{C_{ag}} - \frac{eN_b}{C_{bg}} \end{pmatrix} \ . \tag{J.4}$$

The linear algebra of this particular problem turns out to be rather complicated and is best calculated with a software package such as Mathematica. For $C_l = C_{ab}$, the

solution of this matrix equation is that of equation (8)-(10) in Ref. [89][2]. Those formulas were given in eq. 3.63.

We are interested what effect the coupling capacitance has on the charging energy. It is clear that the coupling to the leads has some influence on the energy scales within the system, as it has for the single dot in the form of a lever arm. To get some qualitative information about how the system behaves, it is interesting to consider the case $C_l \to 0$. This should result in an expression identical to eq. 3.67. This would be equivalent to solving the linear system

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & \frac{1}{C_{ab}} + \frac{1}{C_{ag}} + \frac{1}{C_{bg}} \end{pmatrix} \begin{pmatrix} q_{ag} \\ q_{bg} \\ q_{ab} \end{pmatrix} = \begin{pmatrix} eN_a \\ eN_b \\ \frac{eN_a}{C_{bg}} - \frac{eN_b}{C_{ag}} \end{pmatrix} . \tag{J.5}$$

Solving eq. J.4 with $C_l \to 0$ yields the same results as solving eq. J.5, as expected. The solution is

$$U(N_a, N_b) = \frac{e^2}{2} \frac{C_{ab}N_a^2 + C_{bg}N_a^2 + 2C_{ab}N_aN_b + C_{ab}N_b^2 + C_{ag}N_b^2}{C_{ab}C_{ag} + C_{ab}C_{bg} + C_{ag}C_{bg}} - e(N_a + N_b)V_g \, . \tag{J.6}$$

For tunneling to take place, we must fulfill the conditions

$$\begin{aligned} U(N_a, N_b) &= U(N_a - 1, N_b) \\ U(N_a, N_b) &= U(N_a, N_b - 1) \, . \end{aligned} \tag{J.7}$$

This leads to the conditions:

$$\begin{aligned} \frac{V_g}{e} &= \frac{(C_{ab} + C_{bg})(N_a - \frac{1}{2}) + C_{ab}N_b}{C_{ab}C_{ag} + C_{ab}C_{bg} + C_{ag}C_{bg}} \\ \frac{V_g}{e} &= \frac{(C_{ab} + C_{ag})(N_b - \frac{1}{2}) + C_{ab}N_a}{C_{ab}C_{ag} + C_{ab}C_{bg} + C_{ag}C_{bg}} \end{aligned} \tag{J.8}$$

[2]We again mention that there is a mistake in formula (9) in I. M. Ruzin's paper. The index C_i should read $C_{\neg i}$, meaning "not i".

These conditions have the right limits, so that in the independent dot limit, $C_{ab} \ll C_{ag},\ C_{bg}$

$$\begin{aligned} V_g &= \frac{e}{C_{ag}}(N_a - \frac{1}{2}) \\ V_g &= \frac{e}{C_{bg}}(N_b - \frac{1}{2}) \ . \end{aligned} \tag{J.9}$$

and in the single dot limit $C_{ab} \gg C_{ag},\ C_{bg}$

$$V_g = \frac{e}{C_{ag} + C_{bg}}(N_a + N_b - \frac{1}{2}) \ . \tag{J.10}$$

This means that for vanishing values of the coupling capacitance we expect conductance peaks only if both conditions J.9 are fulfilled, yielding random peaks.

Bibliography

[1] K. R. Early, *Experimental Characterization and Physical Modeling of Resolution Limits in Proximity Printing X-Ray Lithography.* PhD thesis, Massachusetts Institute of Technology, Aug. 1991.

[2] K. Ismail, W. Chu, D. A. Antoniadis, and H. I. Smith, "Surface-superlattice effects in a grating-gate GaAs/GaAlAs modulation doped field-effect transistor," *Appl. Phys. Lett.*, vol. 52, no. 13, pp. 1071–1073, Mar. 1988.

[3] K. Ismail, W. Chu, D. A. Antoniadis, and H. I. Smith, "Lateral-surface-superlattice and quasi-one-dimensional GaAs/GaAlAs modulation-doped field-effect transistors fabricated using x-ray and deep-ultraviolet lithography," *J. Vac. Sci. Technol. B*, vol. 6, no. 6, pp. 1824–1827, Nov/Dec 1988.

[4] K. Ismail, *The Study of Electron Transport in Field-Effect-Induced Quantum wells on GaAs/GaAlAs.* PhD thesis, Massachusetts Institute of Technology, May 1989.

[5] K. Ismail, "Resonant tunneling across and mobility modulation along surface-structured quantum wells," *J. Vac. Sci. Technol. B*, vol. 7, no. 6, pp. 2025–2029, Nov/Dec 1989.

[6] K. Ismail, D. A. Antoniadis, and H. I. Smith, "Lateral resonant tunneling in a double-barrier field-effect transistor," *Appl. Phys. Lett.*, vol. 55, no. 6, pp. 589–591, Aug. 1989.

[7] K. Ismail, W. Chu, A. Yen, D. A. Antoniadis, and H. I. Smith, "Negative transconductance and negative differential resistance in a grid-gate modulation-doped field-effect transistor," *Appl. Phys. Lett.*, vol. 54, no. 5, pp. 460–462, Jan. 1989.

[8] A. Toriumi, K. Ismail, M. Burkhardt, D. A. Antoniadis, and H. I. Smith, "Resonant magnetoconductance in a two-dimensional lateral-surface-superlattice," *Phys. Rev. B*, vol. 41, no. 17, pp. 12346–12349, 1990.

[9] R. A. Ghanbari, W. Chu, E. E. Moon, M. Burkhardt, K. Yee, D. A. Antoniadis, H. I. Smith, M. L. Schattenburg, K. W. Rhee, R. Bass, M. C. Peckerar, and

M. R. Melloch, "Fabrication of parallel quasi-one-dimensional wires using a novel conformable x-ray mask technology," *J. Vac. Sci. Technol. B*, vol. 10, no. 6, pp. 3196–3199, 1992.

[10] M. Burkhardt, D. A. Antoniadis, T. P. Orlando, H. I. Smith, M. R. Melloch, K. W. Rhee, and M. C. Peckerar, "Fabrication using x-ray nanolithography and measurement of coulomb blockade in a variable-sized quantum dot," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, pp. 3611–3613, Nov/Dec 1994.

[11] I. Y. Yang, H. Hu, L. T. Su, V. V. Wong, M. Burkhardt, E. E. Moon, J. M. Carter, D. A. Antoniadis, H. I. Smith, K. W. Rhee, and W. Chu, "High performance self-aligned sub-100 nm metal-oxide semiconductor field-effect transistors using x-ray lithography," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, pp. 4051–4054, Nov/Dec 1994.

[12] G. G. Shahidi, "Electron velocity overshoot at room and liquid nitrogen temperatures in silicon inversion layers," *IEEE Trans. Electron Devices*, vol. 9, no. 2, pp. 94–96, Feb. 1988.

[13] N. Gupta, S. D. Hector, K. W. Rhee, and H. I. Smith, "Fabrication of 100 nm T-gates for monolithic microwave integrated circuits using x-ray lithography," *J. Vac. Sci. Technol. B*, vol. 11, no. 6, pp. 2625–2628, Nov/Dec 1993.

[14] C. T. Liu, D. C. Tsui, M. Shayegan, K. Ismail, M. Burkhardt, D. A. Antoniadis, and H. I. Smith, "Observation of Landau level splitting in two-dimensional lateral surface superlattices," in *The Physics of Semiconductors*, vol. 2, pp. 1701–1704, World Scientific, 1990.

[15] K. Ismail, M. Burkhardt, H. I. Smith, N. H. Karam, and P. A. Sekula-Moise, "Patterning and characterization of large-area quantum wire arrays," *Appl. Phys. Lett.*, vol. 58, no. 14, pp. 1539–1541, Apr. 1991.

[16] V. V. Wong, J. Ferrera, J. N. Damask, J. M. Carter, E. E. Moon, H. A. Haus, H. I. Smith, and S. Rishton, "Spatial-phase-locked electron-beam lithography and x-ray lithography for fabricating first-order gratings on rib waveguides," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, pp. 3741–3745, Nov/Dec 1994.

[17] H. I. Smith and D. A. Antoniadis, "Seeking a radically new electronics," *Technology Review*, vol. 93, p. 26, Apr. 1990.

[18] S. D. Hector, M. L. Schattenburg, E. H. Anderson, W. Chu, V. V. Wong, and H. I. Smith, "Modeling and experimental verification of illumination and diffraction effects on image quality in x-ray lithography," *J. Vac. Sci. Technol. B*, vol. 10, no. 6, pp. 3164–3168, 1992.

[19] S. D. Hector, V. V. Wong, H. I. Smith, M. A. McCord, and K. W. Rhee, "Printability of sub-150 nm features in x-ray lithography: Theory and experiments," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, pp. 3965–3969, Nov/Dec 1994.

[20] S. D. Hector, *Optimization of Image Formation in X-Ray Lithography Using Rigorous Electromagnetic Theory and Experiments.* PhD thesis, Massachusetts Institute of Technology, May 1994.

[21] W. Chu, H. I. Smith, and M. L. Schattenburg, "Replication of 50-nm-linewidth device patterns using proximity x-ray lithography at large gaps," *Appl. Phys. Lett.*, vol. 59, no. 13, pp. 1641–1643, Sept. 1991.

[22] Y. Chen, R. Kupka, F. Rousseaux, F. Carcenac, D. Decanini, M. F. Ravet, and H. Launois, "50-nm x-ray lithography using synchrotoron radiation," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, pp. 3959–3964, Nov/Dec 1994.

[23] R. Sherman, J. Grob, and W. Whitlock, "Dry surface cleaning using CO_2 snow," *J. Vac. Sci. Technol. B*, vol. 9, no. 4, p. 1970, Jul/Aug 1991.

[24] W. Chu, *Inorganic X-ray Mask Technology for Quantum-Effect Devices.* PhD thesis, Massachusetts Institute of Technology, Feb. 1993.

[25] Y.-C. Ku, *Fabrication of Distortion Free X-ray Masks Using Low Stress Tungsten.* PhD thesis, Massachusetts Institute of Technology, October 1991.

[26] A. Moel, W. Chu, K. Early, Y.-C. Ku, E. E. Moon, F. Tsai, H. I. Smith, M. L. Schattenburg, C. D. Fung, F. W. Griffith, and L. E. Haas, "Fabrication and characterization of high-flatness mesa-etched silicon nitride x-ray masks," *J. Vac. Sci. Technol. B*, vol. 9, pp. 3287–3291, 1991.

[27] I. Y. Yang, J. M. Carter, S. Silverman, S. Rishton, D. A. Antoniadis, and H. I. Smith, "Combining and matching optical, e-beam and x-ray lithographies in the fabrication of Si CMOS circuits with 0.1 and sub-0.1 μm features." submitted to EIPB'95, Scottsdale, Az.

[28] M. L. Schattenburg, N. A. Polce, H. I. Smith, and R. Stein, "Fabrication of flip-bonded mesa masks for x-ray lithography," *J. Vac. Sci. Technol. B*, vol. 11, no. 6, pp. 2906–2909, Nov/Dec 1993.

[29] W. Chu, M. L. Schattenburg, and H. I. Smith, "Low-stress gold electroplating for x-ray masks," in *Microelectronic Engineering*, vol. 17, (North-Holland), pp. 223–226, Elsevier Science Publishers B. V., 1992.

[30] W. J. Dauksher, D. J. Resnick, W. A. Johnson, and A. W. Yanof, "A new operating regime for electroplating the gold absorber on x-ray masks," in *Microelectronic Engineering*, vol. 23, pp. 235–238, Elsevier Science Publishers B. V., 1994.

[31] W. J. Dauksher, D. J. Resnick, P. A. Seese, K. D. Cummings, A. W. Yanof, and W. A. Johnson, "Effect of brightener concentration on the thermal distortion of gold plated x-ray masks," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, pp. 3990–3994, Nov/Dec 1994.

[32] R. E. Acosta, W. A. Johnson, B. S. Berry, and W. C. Pritchet, "Annealing behavior of gold absorber in x-ray masks," in *Microelectronic Engineering*, vol. 17, (North-Holland), pp. 189–192, Elsevier Science Publishers B. V., 1992.

[33] Y.-C. Ku, L. P. Ng, R. Carpenter, K. Lu, H. I. Smith, L. E. Haas, and I. Plotnik, "In-situ stress monitoring and deposition of zero stress w for x-ray masks," *J. Vac. Sci. Technol. B*, vol. 9, p. 3297, 1991.

[34] M. Mondol, H. Li, G. Owen, and H. I. Smith, "Uniform-stress tungsten on x-ray mask membranes via He-backside temperature homogenization," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, pp. 4024–4027, Nov/Dec 1994.

[35] M. L. Schattenburg, I. Tanaka, and H. I. Smith, "Microgap x-ray nanolithography," in *Microelectronic Engineering*, vol. 6, (North-Holland), pp. 273–279, Elsevier Science Publishers B. V., 1987.

[36] W. Chu, C. C. Eugster, A. Moel, E. E. Moon, M. L. Schattenburg, H. I. Smith, K. W. Rhee, M. C. Peckerar, and M. R. Melloch, "Conductance quantization in a GaAs electron waveguide device fabricated by x-ray lithography," *J. Vac. Sci. Technol. B*, vol. 10, no. 6, pp. 2966–1969, 1992.

[37] R. A. Ghanbari, *Physics and Fabrication of Quasi-One-Dimensional Conductors.* PhD thesis, Massachusetts Institute of Technology, Feb. 1993.

[38] M. Burkhardt, H. I. Smith, D. A. Antoniadis, T. P. Orlando, M. R. Melloch, K. W. Rhee, and M. C. Peckerar, "Gap control in the fabrication of quantum-effect devices using x-ray nanolithography," in *Micro- and Nano-Engineering*, 1994. to be published.

[39] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, vol. 7 of *Course of Theoretical Physics.* Oxford: Pergamon Press, third ed., 1986. §14, see Problem 2.

[40] A. Kumar, C. C. Eugster, T. P. Orlando, D. A. Antoniadis, J. M. Kinaret, M. J. Rooks, and M. Melloch, "Correlation of oscillations in a quantum dot with three contacts." submitted to Appl. Phys. Lett., 1994.

[41] A. Kumar, J. Kinaret, C. C. Eugster, T. P. Orlando, D. A. Antoniadis, M. J. Rooks, and M. Melloch, "Anti-correlated oscillations in a three-lead quantum dot." Villars-sur-Ollon, Switzerland, 22-29 January 1994 (to be published in

the proceedings). presented at *Rencontres de Moriond: Coulomb Blockade and Interference Effects in Small Electronic Structures.*

[42] L. T. Romankiw. private communication. The brightener is supposedly a common chemical. It is safe, cheap, and most likely available in an MIT laboratory supply.

[43] R. Dingle, M. D. Feuer, and C. W. Tu, *The Selectively Doped Heterostructure Transistor: Materials, Devices, and Circuits*, vol. 11 of *VLSI Electronics Microstructure Science*, ch. 6, pp. 215–264. Academic Press, 1985.

[44] M. J. Rooks, C. C. Eugster, J. A. del Alamo, G. L. Snider, and E. L. Hu, "Split-gate electron waveguide fabrication using multilayer poly(methylmethacrylate)," *J. Vac. Sci. Technol. B*, vol. 9, no. 6, pp. 2856–2860, November/December 1991.

[45] R. Ghanbari, M. Burkhardt, D. A. Antoniadis, H. I. Smith, M. R. Melloch, K. W. Rhee, and M. C. Peckerar, "Comparative mobility degradation in modulation-doped GaAs devices after e-beam and x-ray irradiation," *J. Vac. Sci. Technol. B*, vol. 10, no. 6, pp. 2890–2892, 1992.

[46] J. N. Randall, A. C. Seabaugh, and J. H. Luscombe, "Fabrication of lateral resonant tunneling devices," *J. Vac. Sci. Technol. B*, vol. 10, no. 6, pp. 2941–2944, Nov/Dec 1992.

[47] H. I. Smith, S. D. Hector, M. L. Schattenburg, and E. H. Anderson, "A new approach to high fidelity e-beam and ion-beam lithography based on an *in situ* global-fiducial grid," *J. Vac. Sci. Technol. B*, vol. 9, no. 6, pp. 2992–2995, Nov/Dec 1991.

[48] R. Kleinhenz, P. M. Mooney, C. P. Schneider, and O. Paz, "Defects produced in silicon and GaAs during e-beam evaporation of metals," in *13th International Conference on Defects in Semiconductors*, vol. 14a, the Metallurgical Society of AIME, Warrendale, PA, 1985, 1985.

[49] M. Nel and F. D. Auret, "Deep-level transient spectroscopy detection of defects created in epitaxial GaAs after electron-beam metallization," *J. Appl. Phys.*, vol. 64, no. 5, pp. 2422–2425, Sept. 1988.

[50] M. Nel and F. D. Auret, "Defect formation in GaAs by subthreshold energy (0.2-3 keV) electron irradiation," *Jpn. J. Appl. Phys.*, vol. 28, no. 12, pp. 2430–2435, Dec. 1989.

[51] T. H. Ning, "Electron trapping in SiO_2 due to electron-beam deposition of aluminum," *J. Appl. Phys.*, vol. 49, no. 7, pp. 4077–4082, July 1978.

[52] D. D. Smith, T. Fink, W. D. Braddock, and M.-L. Saunders, "RTA removal of e-beam-induced damage in GaAs-AlGaAs heterostructures as determined by magnetotransport characterization," *Journal of Electronic Materials*, vol. 19, no. 7, pp. 757–760, 1990.

[53] L. Pfeiffer, K. W. West, H. L. Störmer, and K. W. Baldwin, "Electron mobilities exceeding 10^7 cm^2/Vs in modulation-doped GaAs," *Appl. Phys. Lett.*, vol. 55, pp. 1888–1890, 1988.

[54] D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, "One-dimensional transport and the quantization of the ballistic resistance," *J. Phys. C: Solid State Phys.*, vol. 21, pp. L209–L214, 1988.

[55] D. A. Wharam, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, "Addition of the one-dimensional quantized ballistic resistance," *J. Phys. C: Solid State Phys.*, vol. 21, pp. L887–L891, 1988.

[56] L. P. Kouwenhoven, B. J. van Wees, C. P. M. Harmans, J. G. Williamson, H. van Houten, C. W. J. Beenakker, C. T. Foxon, and J. J. Harris, "Nonlinear conductance of quantum point contacts," *Phys. Rev. B*, vol. 39, no. 11, pp. 8040–8043, Apr. 1989.

[57] P. F. Bagwell and T. P. Orlando, "Landauer's conductance formula and its generalization to finite voltages," *Phys. Rev. B*, vol. 40, no. 3, pp. 1456–1464, July 1989.

[58] G. L. Timp and R. E. Howard, "Quantum mechanical aspects of transport in nanoelectronics," *IEEE Proceedings*, vol. 79, no. 8, pp. 1188–1207, Aug. 1991.

[59] C. C. Eugster, *Electron Waveguide Devices.* PhD thesis, Massachusetts Institute of Technology, May 1993.

[60] J. H. Davies, "Electronic states in narrow semiconducting wires near threshold," *Semiconductur Science and Technology*, vol. 3, pp. 995–1009, Oct. 1988.

[61] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM J. Res. Develop.*, vol. 1, pp. 223–231, July 1957.

[62] R. Landauer, "Can a length of perfect conductor have a resistance?," *Phys. Lett.*, vol. 85A, no. 2, pp. 91–93, Sept. 1981.

[63] R. Landauer, "Conductance determined by transmission: Probes and quantised constriction resistance," *J. Phys.: Condens. Matter*, vol. 1, pp. 8099–8110, 1989.

[64] J. H. F. Scott-Thomas, M. A. Kastner, D. A. Antoniadis, H. I. Smith, and S. Field, "Si metal-oxide semiconductor field effect transistor with 70-nm slotted gates for study of quasi-one-dimensional quantum transport," *J. Vac. Sci. Technol. B*, vol. 6, no. 6, pp. 1841–1844, Nov/Dec 1988.

[65] J. H. F. Scott-Thomas, S. B. Field, M. A. Kastner, H. I. Smith, and D. A. Antoniadis, "Conductance oscillations periodic in the density of a one-dimensional electron gas," *Phys. Rev. Lett.*, vol. 62, no. 5, pp. 583–586, 1989.

[66] H. van Houten and C. W. J. Beenakker, "Comment on *conductance oscillations periodic in the density of a one-dimensional electron gas*," *Phys. Rev. Lett.*, vol. 63, no. 17, p. 1893, 1989.

[67] L. I. Glazman and R. I. Shekhter, "Coulomb oscillations of the conductance in a laterally confined heterostructure," *J. Phys.: Condens. Matter*, vol. 1, pp. 5811–5815, 1989. This article was submitted after Scott-Thomas's measurements were published but before the explanation by van Houten. It does not cite experimental results.

[68] K. K. Likharev, "Correlated discrete transfer of single electrons in ultrasmall tunnel junctions," *IBM J. Res. Develop.*, vol. 32, no. 1, pp. 144–158, Jan. 1988.

[69] E. B. Foxman, *Single Electron Charging and Quantum Effects in Semiconductor Nanostructures.* PhD thesis, Massachusetts Institute of Technology, Aug. 1993.

[70] L. I. Glazman and R. I. Shekhter, "Coulomb oscillations of the conductance in a laterally confined heterostructure," *J. Phys.: Condens. Matter*, vol. 1, pp. 5811–5815, 1989.

[71] U. Meirav, M. A. Kastner, and S. J. Wind, "Single-electron charging and periodic conductance resonances in GaAs nanostructures," *Phys. Rev. Lett.*, vol. 65, no. 6, pp. 771–774, 1990.

[72] L. P. Kouwenhoven, N. C. van der Vaart, A. T. Johnson, W. Kool, C. J. P. M. Harmans, J. G. Williamson, A. A. M. Staring, and C. T. Foxon, "Single electron charging effects in semiconductor quantum dots," *Zeitschrift für Physik B – Condensed Matter*, vol. 85, pp. 367–373, 1991.

[73] L. P. Kouwenhoven, *Transport of Electron-Waves and Single-Charges in Semiconductor Nanostructures.* PhD thesis, Delft University of Technology, 1992.

[74] A. A. M. Staring, H. van Houten, C. W. J. Beenakker, and C. T. Foxon, *High Magnetic Fields in Semiconductor Physics III*, ch. Coulomb-Regulated Conductance Oscillations in a Disordered Quantum Wire, pp. 301–312. Springer-Verlag, 1992.

[75] H. van Houten, C. W. J. Beenakker, and A. A. M. Staring, "Coulomb-blockade oscillations in quantum wires and dots," in *Single-Electron Tunneling and Mesoscopic Devices*, pp. 159–169, Springer Verlag, 1992.

[76] H. van Houten, C. W. J. Beenakker, and A. A. M. Staring, *Coulomb-Blockade Oscillations in Semiconductor Nanostructures*, ch. 5, pp. 167–216. Plenum Press, 1992.

[77] M. A. Kastner, "Artificial atoms," *Physics Today*, pp. 24–31, Jan. 1993.

[78] A. Kumar, *Single Electron Charging Effects in Quantum Dot Nanostructures.* PhD thesis, Massachusetts Institute of Technology, 1994.

[79] L. P. Kouwenhoven, A. T. Johnson, N. C. van der Vaart, C. J. P. M. Harmans, and C. T. Foxon, "Quantized current in a quantum-dot turnstile using oscillating tunnel barriers," *Phys. Rev. Lett.*, vol. 67, no. 12, pp. 1626–1629, Sept. 1991.

[80] J. R. Tucker, "Complementary digital logic based on the "Coulomb blockade"," *J. Appl. Phys.*, vol. 72, no. 9, pp. 4399–4413, Nov. 1992.

[81] K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai, and K. Seki, "Room-temperature single-electron memory," *IEEE Trans. Electron Devices*, vol. 41, no. 9, pp. 1628–1638, Sept. 1994.

[82] D. J. Paul, J. R. A. Cleaver, H. Ahmed, and T. E. Whall, "Coulomb blockade in silicon based structures at temperatures up to 50 K," *Appl. Phys. Lett.*, vol. 63, no. 5, pp. 631–632, Aug. 1993.

[83] Y. Takahashi, M. Nagase, H. Namatsu, K. Kurihara, K. Iwdate, Y. Nakajima, S. Horiguchi, K. Murase, and M. Tabe, "Conductance oscillations of a si single electron transistor at room temperature," Dec. 1994. IEDM'94.

[84] L. D. Landau and E. M. Lifschitz, *Statistische Physik*, vol. 5 of *Lehrbuch der theoretischen Physik.* Berlin: Akademie-Verlag, 7th ed., 1987. translated from russian.

[85] C. W. J. Beenakker, "Theory of Coulomb-blockade oscillations in the conductance of a quantum dot," *Phys. Rev. B*, vol. 44, no. 4, pp. 1646–1656, July 1991.

[86] H. A. Haus and J. R. Melcher, *Electromagnetic Fields and Energy.* Englewood Cliffs, NJ: Prentice Hall, 1989.

[87] I. O. Kulik and R. I. Shekhter, "Kinetic-phenomena and charge discreteness in granular media," *JETP*, vol. 68, p. 623, 1975.

[88] J. Hugunin. private communication. The sample temperature might be higher than the sample block temperature; instead of 300 mK, it might be 800 mK.

[89] I. M. Ruzin, V. Chandrasekhar, E. I. Levin, and L. I. Glazman, "Stochastic coulomb blockade in double-dot system," *Phys. Rev. B*, vol. 45, no. 23, pp. 13469–13478, June 1992.

[90] L. I. Glazman and V. Chandrasekhar, "Coulomb blockade oscillations in a double-dot system," *Europhysics Letters*, vol. 19, no. 7, pp. 623–628, Aug. 1992.

[91] L. I. Glazman and I. M. Ruzin, "Metal to insulator crossover in mesoscopic wires," *Physica Scripta*, vol. T42, pp. 122–132, 1992.

[92] G. Klimeck, G. Chen, and S. Datta, "Conductance spectroscopy in coupled quantum dots," *Phys. Rev. B*, vol. 50, no. 4, pp. 2316–2324, July 1994.

[93] G. Chen, G. Klimeck, S. Datta, G. Chen, and W. A. Goddard III, "Resonant tunneling through quantum-dot arrays," *Phys. Rev. B*, vol. 50, no. 11, pp. 8035–8038, sep 1994.

[94] F. R. Waugh, M. J. Berry, D. J. Mar, R. M. Westervelt, K. C. Campman, and A. C. Gossard, "Single-electron charging in double and triple quantum dots with tuneable coupling," 1994. submitted to Phys. Rev. Lett.

[95] M. Kemerink and . M. L. W, "Stochastic Coulomb blockade in a double quantum dot," *Appl. Phys. Lett.*, vol. 65, no. 8, pp. 1012–1014, Aug. 1994.

[96] A. Toriumi, K. Ismail, M. Burkhardt, D. A. Antoniadis, and H. I. Smith, "Resonant magneto-conductance in a two-dimensional lateral-surface-superlattice," in *The Physics of Semiconductors*, vol. 2, pp. 1313–1316, World Scientific, 1990.

[97] U. Tietze and C. Schenk, *Halbleiterschaltungstechnik.* Berlin: Springer-Verlag, 6th ed., 1983. english translation exists.

[98] P. Horowitz and W. Hill, *The Art of Electronics.* Cambridge: Cambridge University Press, 2nd ed., 1989.

[99] A. Moel, E. E. Moon, R. Frankel, and H. I. Smith, "Novel on-axis interferometric alignment metod with sub-10 nm precision," *J. Vac. Sci. Technol. B*, vol. 11, no. 6, pp. 2191–2194, Nov/Dec 1993.

[100] B. L. Henke and M. A. Tester, "Techniques of low energy x-ray spectroscopy (0.1 to 2 keV region)," in *Advances in X-ray Analysis*, vol. 18, Plenum Press, 1975.

[101] S. D. Hector, H. I. Smith, and M. L. Schattenburg, "Simultaneous optimization of spectrum, spatial coherence, gap, feature bias, and absorber thickness in synchrotron-based x-ray lithography," *J. Vac. Sci. Technol. B*, vol. 11, no. 6, pp. 2981–2985, 1993.

[102] H. W. Ott, *Noise Reduction Techniques in Electronic Systems.* New York: John Wiley & Sons, 2nd ed., 1988.